Langley

CELL
FUNCTION

Plate 1. Types of Cells Found in Normal Blood. Colored by Wright's stain. All cells drawn with the same magnification: × 1150. Numbers 1-5, neutrophils; 6-8, lymphocytes; 9-13, monocytes; 14, platelets; 15, erythrocytes; 16, 17, eosinophils; 18-20, "basket cells" (degenerated leukocytes); 21, basophilic leukocyte.

Cell Function

REINHOLD BOOKS IN THE BIOLOGICAL SCIENCES

CONSULTING EDITOR'S STATEMENT

MANY WRITERS on cell physiology seem unable to make up their minds whether it is the cell, or the physiology, with which they should be preoccupied. Dr. L. L. Langley is far too experienced, both as an author and as a teacher, to fall into this common error. He has produced, in the present volume, one of the best balanced books yet to be written about this field. This balance is not only between cellular structure and function but also between the contributions which the cell makes to the structure and function of the whole organism. Thus the reader is led from the functional anatomy of the cell to the study of protoplasm, and thence to the cell membrane which integrates the cell's internal and external environment. Next, intracellular activities are discussed in detail, leaving the third section of the book to discuss organ systems from the viewpoint of a cell physiologist.

None of this can be done, of course, unless the student has some basic understanding of the physical sciences, but even the student who lacks this, need not turn from the present volume, since the last four chapters— by no means to be referred to as an appendix—survey in clear language the knowledge expected of the reader. I like Dr. Langley's idea of placing this material at the end where it is available to those who want it without getting in the way of those competent to plunge directly into the heart of the matter.

I feel that this addition to the REINHOLD BOOKS IN THE BIOLOGICAL SCIENCES not only fills a long-felt want but fills it extremely well.

PETER GRAY

Pittsburgh, Pennsylvania
April, 1961

CELL
FUNCTION

An Introduction to the Physiology of the
Cell and Its Role in the Intact Organism

L. L. LANGLEY, Ph.D., LL.B.

Professor of Physiology,
University of Alabama

ILLUSTRATED BY FRANCES LANGLEY

PREFACE

COURSES OFFERED under the title of "Cell Physiology" are as varied
as are cells themselves. The indecision as to exact boundaries and
content of this subject is underscored by the apparent difficulty
teachers have in deciding what to call it. Is it Cell Physiology, or is it
General Physiology? Some semanticists insist that the term "Cell
Physiology" is too narrow, that there are forms of life, such as filter-
able viruses, that cannot truly be called cells. They go on to insist
that it is necessary, in many instances, to study masses of cells, for
example, the liver, when considering metabolism. Thus, they prefer
to call their field, and the course they teach, "General Physiology."
Theoretically, they cut a broad path through all of the biological
subdivisions extracting principles that are common to all. This is a
bold concept, a very large undertaking. Nothing so ambitious has
been attempted here. This textbook is but an introduction, an in-
troduction to merely a few of the basic principles governing living
processes.

The cell is the fundamental unit, the building block, so to speak,
of almost all living forms. Accordingly, the functional anatomy of
the generalized cell and basic intracellular activities are considered.
But, save for organisms such as the unicellular ameba, cells do not
function as independent entities. They are united into tissues, into
glands, into multicellular organisms. Among the most highly de-
veloped of these are the mammals. Thus it seems logical to examine
how cells are united into tissues, into complex structures and, finally,
to survey mammalian physiology. Accordingly, although this is an
introductory survey, it is, nonetheless, comprehensive in the sense

that not only are intracellular processes considered but also the interrelationships between cells as exemplified in the most physiologically advanced organisms.

Physiology is a study that depends upon many other sciences. How can metabolism be understood if the student does not have a knowledge of chemistry? How can the propagation of an impulse, or the circulation of blood be comprehended without a knowledge of physics? How can fact be differentiated from fancy without an appreciation of the scientific method? And how can data, the essence of any science, be handled without a grounding in mathematics? Since the answers to these questions are self-evident, a section devoted to basic background principles has been included. It is presumed that students using this textbook have had courses in chemistry, physics, and mathematics, but it is also highly probable that certain principles in these disciplines may have been missed or forgotten. Accordingly, Part IV of this textbook is available for reference. In short, the necessary tools are there; if the student already has his own equipment he will not need them, but if he has forgotten any of them, they are provided for his use.

If I have learned anything as author, or co-author, of five other textbooks, it is that a text free of errors is a shimmering Goddess, an unattainable ideal. But that goal can be approached with the assistance of others, with the assistance of dedicated experts in the field. I have been fortunate in having three such experts comb the manuscript. They are Dr. Herbert S. Poland, Dr. John Spikes, and Dr. William H. Johnson. I also owe deep appreciation to Mr. John Hart for his careful editorial assistance during the final preparation of the manuscript. These individuals have contributed greatly, but whatever shortcomings still remain are, of course, mine alone.

Unlike wine, a textbook is not enhanced by quiet, unused, undisturbed aging. A textbook improves only if it is vigorously used, mercilessly criticized. If this book is to survive, to reach vintage caliber, it will be because the readers have been free with their comments.

L. L. LANGLEY

Birmingham, Alabama
February, 1961

CONTENTS

NOTE: The author is indebted to Professor Sir Ronald A. Fisher, F.R.S.,
Cambridge, and to Dr. Frank Yates, F.R.S., Rothamsted; also to Messrs.
Oliver & Boyd Ltd., Edinburgh, for permission to reprint Tables Nos.
III and IV from their book, "Statistical Tables for Biological, Agricultural
and Medical Research."

PART I

Functional
Anatomy
of the Cell

ORGANIZATION

THE BASIC UNIT of all living organisms is the cell. Some forms, such as the ameba, consist of but a single cell. In the more complex forms there are millions of these units; units that have been differentiated into a remarkable variety of types to form the various tissues of the body. Although there is this great differentiation, a fundamental pattern of organization persists.

With the technical tools available today one would think that the anatomy of a typical cell would be known thoroughly and without question. Such is far from the fact. The structure of the cell is still the subject of a heated debate which newer methods of investigation have only intensified. It is beyond the scope of this text to discuss all of the conflicting evidence but an indication will be given where conflicting views exist.

THE GENERALIZED CELL

Although there is a wide variety of cells, all of them have many features in common. Figure 1.1, page 6, shows a generalized cell. It is seen to contain a prominent nucleus. In addition to the nucleus most cells contain structures termed: 1) **organelles**, and 2) **inclusions.** An organelle is a structure with definite functions. An inclusion is a structure which is thought not to have a specific function, or to have one that is not essential to cell survival. Organelles include: 1) the **centrosome,** 2) the **mitochondria,** and 3) the **Golgi apparatus.** Inclusions are usually secretory granules, pigments, and stored foods, such

3

as fat depots. The entire cell is generally surrounded and contained by a membrane.

CYTOPLASM

The terms cytoplasm and protoplasm are often used synonymously. It is better, however, to reserve the term protoplasm to cover the general sum total of cellular contents which as a whole exhibits the properties associated with a living system. Cytoplasm, then, refers to the protoplasm that lies between the nucleus and the cell membrane. Within the nucleus also there is protoplasm, termed nucleoplasm.

Cytoplasm has been described as having an "optical emptiness." Such a description is indicative of a conception of cellular structure that relegates many well known, and generally accepted cell constituents to the category of artifacts, that is, structures that have been inadvertently produced. However, the better opinion is that cytoplasm is not simply what remains in the cell after various structures have been removed, but rather that the cytoplasm itself possesses structure. All components that exist in the cytoplasm in a constant and ubiquitous fashion should be thought of as essential elements.

By the use of phase and electron microscopy it has become clear that there are not only definite, functional structures (organelles) lying in the cytoplasm, but that the cytoplasm itself has a discernible and significant morphology.

Ergastoplasm

Under the electron microscope a very fine structural network, a reticulum, can be seen in the cytoplasm of almost all cells with the exception of the mature erythrocyte. Because of its reticular appearance (see Plate 2 on facing page) the term **endoplasmic reticulum** has been used to describe this network. However, most authorities still prefer to use the older term, ergastoplasm.

The origin and function of the ergastoplasm remain in question. There is some evidence that it is derived from the nuclear membrane. If this is true, then there is an explanation for the absence of ergastoplasm in mature erythrocytes since they do not contain nuclei. It has also been suggested that mitochondria may have a significant role in ergastoplasm formation. Finally, the possibility exists, as always, that what is described as ergastoplasm is nothing but an artifact.
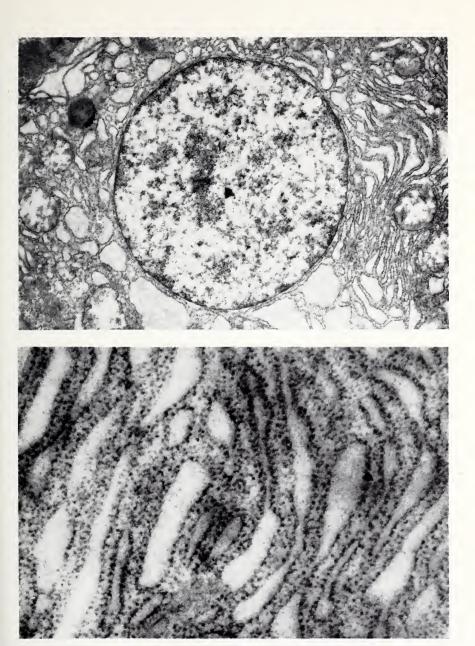
Plate 2. Ergastoplasm. In the upper photomicrograph the relationship of the ergastoplasm to the nucleus is clearly shown. (Rat pancreas, magnification × 24,000, photographic reduction 30%, courtesy of Dr. W. Bernhard, Institut de Recherches sur le Cancer, Paris.) The lower photomicrograph shows the fine detail of the ergastoplasm. (Magnification × 100,000, photographic reduction 30%, courtesy of Dr. G. E. Palade, The Rockefeller Institute for Medical Research, N. Y.)

Microsomes

Small granules, termed microsomes, or **Claude's particles** have been described as lying in the cytoplasm. They range from 60 to 200 milli-microns (micron = one-millionth of a meter) in diameter and therefore can only be visualized with the aid of the electron microscope. Whether they are independent structures, or merely a part of the ergastoplasm has not been settled. And, as in the case of ergastoplasm, they may represent nothing but artifacts. The better concept appears to be that ergastoplasm is not an artifact and that microsomes are fragments of ergastoplasm.

Biochemical study of microsomes discloses that they are composed of lipid, mostly phospholipid, ribonucleic acid, and protein. It is perhaps significant that about 60 percent of all the ribonucleic acid of the entire cell is concentrated in the microsomes.

The function of the microsomes is almost purely a matter of speculation. It has been suggested that the microsomes play an important role in protein synthesis by the cell. They may also be involved in lipid metabolism.

Hyaloplasm

After a cell has been subjected to centrifugation a clear fluid remains which is called the **ground substance,** or hyaloplasm. In other words, the cell is a mass of protoplasm in which there are various structures. When these are swept to one side by centrifugation there remains only a clear, homogeneous fluid called the hyaloplasm. It should be noted that low speed centrifugation causes only the larger structures to be displaced and the ergastoplasm is left undisturbed. However, with very high speeds it is possible to disrupt the ergasto-plasm and to cause the microsomes to be dislodged.

MITOCHONDRIA

The noted cytologist, Jean Brachet, exclaimed: "There is one very good thing about the mitochondria: they certainly exist!" This exclamation vividly reflects the concern, the caution, and the anguish over whether what is seen in the cell is a true structure or an artifact. Quite clearly there is no longer a serious question as to the existence of mitochondria.

Mitochondrial Morphology

Mitochondria, also called **chondriosomes,** usually have a rod shape, although they do sometimes appear as spherical granules (Fig. 1.1). The granules have a diameter of about 2 microns, the rods a length of about 5 microns, although much longer ones have been described. It has been shown that mitochondria change shape during the life cycle of the cell so that many shapes and forms are to be seen.

Mitochondria are enveloped by a membrane. In elongated mitochondria, this membrane has been seen to send branches into the substance of the mitochondria and thus divide the structure into compartments.

Mitochondrial Composition

Mitochondria are composed of protein, fat, phospholipid, and a small amount of ribonucleic acid. In addition there are also present
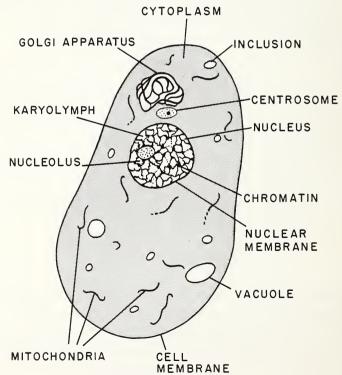
Fig. 1.1. A Generalized Cell.

numerous essential enzymes including cytochrome, succinoxidase, and various phosphatases, as well as enzymes involved in the tricar- boxylic acid cycle (Chapter 5).

Mitochondrial Function

Mitochondria are believed to play an important role in cellular metabolism of oxygen, carbohydrate, fatty acids, and amino acids. Because of the great number of essential enzymes present in the mitochondria and due to the fact that mitochondria apparently enter into all basic metabolic processes, they have been considered to be the "principal power plants of the cell." There is no doubt but that the mitochondrial membrane functions in some manner to regulate this vital role of the mitochondria. It has been shown that mito- chondria may swell or shrink depending upon the type of solution to which they are exposed. The mitochondrial membrane thus must be selectively permeable and control the passage of substances in and out of the mitochondria. The exact nature and function of the mem- brane is not known.

GOLGI APPARATUS

In 1898 Camillo Golgi, an Italian anatomist, described a network of threads in the cytoplasm of a cell that had been stained with silver nitrate or osmic acid. For this discovery, as well as his outstanding investigations of the histology of the nervous system, Golgi was awarded a Nobel prize in 1907.

Unfortunately, in the case of the **Golgi apparatus,** also called Golgi **material** or **complex,** one hesitates to say that it certainly exists, for there is still debate as to whether it is a true constituent of the living cell, or simply an artifact resulting from the method by which the cell is prepared for study. However, the consensus of authorities, at the present time, is that the Golgi apparatus is actually an essential component of the living cell. This conclusion is based on the obser- vation of the Golgi apparatus in the living cell by the use of phase microscopy. Further it is possible to stain a cell with methylene blue without killing it, and under these conditions it is sometimes possible to see the Golgi apparatus.

Golgi Apparatus Morphology

The Golgi apparatus as seen in cells stained with silver nitrate or osmic acid consists of several large, empty vacuoles surrounded by a membrane. There also seems to be present a continuous network of strands, as well as small granules close to the membrane. The entire mass of the Golgi apparatus is relatively large and forms a conspicuous part of the stained cell (Fig. 1.1).

Golgi Apparatus Composition

The Golgi material seems to be quite similar to mitochondria in that it is lipoprotein in structure. Recently, ribonucleic acid (RNA) and alkaline phosphatase have also been identified in the Golgi apparatus.

Golgi Apparatus Function

Due to the fact that the Golgi apparatus is most often present in secretory cells, a secretory function has been attributed to it. It is thought that the Golgi apparatus does not actually take part in the formation of the secretory substance, but rather controls its liberation from the cell. It may be that the Golgi apparatus is some sort of an intracellular pump that regulates the movement of fluids in the cell and the expulsion of secretory products from the cell. In brief, most authorities now believe that the Golgi apparatus truly exists and it appears to be an essential component of the cell.

NUCLEUS

The nucleus is usually spherical in shape and is the most conspicuous structure within the cell (Fig. 1.1). However, in the living cell, not undergoing division, the nucleus is difficult to see unless the phase contrast microscope is used. But after the cell is stained, the nucleus becomes very prominent because it contains **chromatin,** a substance that stains vividly. The chromatin material is essential to heredity. The stained section discloses that the nucleus contains, in addition to the chromatin material, a round structure called a **nucleolus.**

Surrounding the nucleus is a membrane that effectively separates

the cytoplasm of the cell from the nucleus. The protoplasm of the nucleus, **nucleoplasm,** is sometimes referred to as **karyolymph.**

The size of the nucleus varies with that of the cell itself.

Composition of the Nucleus

It is possible to remove the nucleus from a cell. Consequently, nuclear material has been available for detailed study and there is a considerable body of knowledge concerning its chemical composition. The question always arises, however, as to what changes in the chemical composition result from the process of nuclear isolation. Fortunately there are several ways in which nuclear material can be obtained, or it may be analyzed without removal. Thus, the results of various procedures can be compared and used as controls. There is now general agreement as to the main chemical components of the nucleus.

The most important constituent of the nucleus is <u>nucleic acid.</u> Two types of nucleic acid have been identified: 1) **deoxyribonucleic acid (DNA)**, and 2) **ribonucleic acid (RNA)**. In the nucleus there is more DNA than RNA. DNA is a constituent of chromatin. Whether or not DNA is also present in the nucleolus remains uncertain. However, small quantities of DNA have been extracted from the nucleoplasm.

DNA is a complex molecule containing pyrimidines, purines, the sugar deoxyribose, and phosphoric acid. The complexity of the molecule is indicated by the molecular weight which is many millions. Models that have been constructed to depict the DNA molecule show a very long, intertwined, twisted structure (Fig. 1.2). This molecule of DNA is usually closely associated with protein in the nucleus, referred to as **histone,** to form **nucleohistone.** The quantity of nucleohistone remains remarkably constant in the resting cell. During cell division, however, the quantity doubles.

RNA, ribonucleic acid, is found in great amounts in the microsomes and mitochondria. In the nucleus it is most highly concentrated in the nucleolus. It is also present in the chromatin. RNA, like DNA, is a complex molecule which contains pyrimidines, purines, the sugar ribose, and phosphoric acid. The structure of RNA seems to vary with the location in which it is found. That is to say,

RNA of the nucleus is somewhat different in composition from RNA of mitochondria. Too little is known at present concerning the exact structure of RNA to explain these differences. It was pointed out that DNA remains quantitatively constant in the resting cell. This is not true of RNA which shows great quantitative variation from
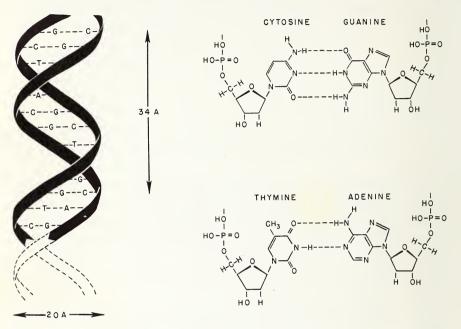
Fig. 1.2.   Schematic Model of a Short Segment of a Molecule of DNA. The two poly-nucleotide chains are shown with the purine and pyrimidine bases represented by the letters A-T (adenine: thymine) and G-C (guanine: cytosine). The specificity of the two chains in the DNA molecule are thought to reside in the hydrogen bonds between the pairs of bases. The details of this hydrogen bonding is shown on the right. The pairs of bases are said to lie in a plane at right angles to the long axis of the molecule. They would represent steps if the structure is compared to a winding staircase. (Redrawn, with permission, from J. Herbert Taylor, American Scientist, 48, 365, 1960.)

cell to cell, and even in the same cell which is subjected to changes in environmental conditions.

The nucleus contains substances other than RNA and DNA. There are present a variety of proteins other than the histones with which DNA combines. In addition, many enzymes are in the nucleus. It is thought significant that the respiratory enzymes, cyto-chrome oxidase and succinic dehydrogenase which are present in

mitochondria, do not exist in the nucleus. On the other hand, many glycolytic enzymes have been isolated and there are also enzymes essential to nucleotide metabolism. The nucleus also contains electrolytes but whether in the same concentration as in the cytoplasm is not known.

Functional Anatomy of the Nucleus

The nucleus is essential to the cell. When it is removed, the cell soon dies. If the nucleus is put into another cell from which its own nucleus has been removed, that cell survives, and if it divides, the new cells usually have characteristics associated with both the mother cell and the new nucleus. In other words, such experiments indicate that the nucleus is not only essential for survival, but that it plays a fundamental role in heredity. This function is considered in detail in Chapter 8.

The high concentration of RNA in the nucleus underscores the conclusion that this structure has a major role in protein synthesis. That RNA is intimately involved in protein synthesis has been inferred from several lines of evidence. In the first place, it has been demonstrated that the RNA concentration is always high in rapidly growing cells. Secondly, in glands in which enzymes are being produced the concentration of RNA is high. Thirdly, cells involved in activities that do not include protein synthesis contain very small quantities of RNA. These observations do not constitute absolute proof, but the fact that the RNA content of various cells always bears such a positive correlation with the degree of protein synthesis certainly makes this conclusion highly probable. Further support is gained from the observation that when a cell changes from a state of low protein synthesis to one of great protein synthesis, the quantity of RNA alters proportionately, and, most significantly, the RNA accumulation occurs before the protein synthesis increases. In addition, it has been shown that ribonuclease, an enzyme that inactivates RNA, inhibits protein synthesis. Finally, and perhaps strongest of all, is the fact that cell-free extracts require RNA of two types, soluble and granular, for protein synthesis to occur. Thus, it may be assumed that RNA plays an essential role in protein synthesis, both in the cellular protoplasm and in the nucleus.

Chromatin. Examination of a stained, resting cell discloses that the chromatin material is dispersed in an irregular pattern throughout the nucleus (Fig. 1.1). Two types of chromatin have been recognized: 1) **euchromatin** and 2) **heterochromatin.** The darker, more dense material is the heterochromatin. The **genes** are thought to be in association with the euchromatin. The difference in staining characteristics of these two types of chromatin may be due to a difference in nucleic acid concentration, or to a greater coiling.

During cell division the chromatin material becomes differentiated into discrete structures termed **chromosomes.** The size of chromosomes varies from about 1 micron in very small cells, up to about 2000 microns in the cells of the salivary glands of *Drosophila.* Detailed study of the chromosome discloses that it is usually divided longitudinally into subdivisions called **chromonemata.** The chromosome has distinct beadlike regions called **chromomeres** (Fig. 1.3).
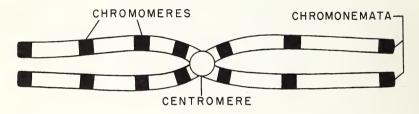
Fig. 1.3.   A Diagrammatic Representation of a Chromosome.

Whether the chromomeres represent true differences in various parts of the chromatin substance, or whether they simply appear different because of coiling of the chromosome has not yet been completely resolved. There is usually present in each chromosome a clear, constricted zone, termed the **centromere.** The centromere has been found to contain less DNA than other parts of the chromosome.

Chromosomes are highly elastic structures. It is possible to stretch them to double their resting length, and when released they return to their original dimensions. This probably reflects the coiled structure that characterizes chromosomes.

The chromomeres give the chromosomes a striped appearance (Fig. 1.3). The dark bands are considered to be disks of DNA. It seems certain that the genes are associated with these dark bands, but whether DNA disks are genes is another question. There is a growing

body of evidence to support the concept that if the DNA disks are not genes, they are certainly essential to the hereditary mechanism. Genes have been shown to exert a catalytic action on the cellular production of enzymes and in this way govern the activity of the cell. DNA could well play such a role.

Nucleolus. The nucleolus is the spherical body generally seen in most nuclei (Fig. 1.1). There may be more than one present. The nucleolus under magnification appears to contain vacuoles and many small granules. The current concept is that the formation of the nucleolus is, in some manner, controlled by the chromatin material, probably by heterochromatin.

There is a high concentration of RNA in the nucleolus and therefore this structure is considered to be of importance in cellular protein synthesis. Usually there is no DNA present.

Nuclear Membrane. The entire nucleus, consisting of the chromatin material, nucleolus, and nuclear protoplasm, is surrounded by a membrane. This membrane may be clearly seen by the use of the electron microscope. It has been found to consist of two layers. The inner layer is continuous and does not appear to contain pores. The outer layer, in contradistinction, does seem to contain pores. Again, however, the charge of artifact has been leveled at these studies, and therefore the question of the actual existence of pores remains unsettled.

Beyond doubt the nuclear membrane is permeable. In fact it is so permeable, that even proteins get through, at least when the nucleus is isolated from the cell. The question, then, is whether or not the process of isolation alters the nuclear-membrane permeability. The current evidence supports the conclusion that even in the intact state proteins with a molecular weight up to about 15,000 penetrate the nuclear membrane.

OTHER CELL CONSTITUENTS

Centrosome

In the resting cell the centrosome appears as a clear area near the nucleus in which there is a very small granule, the centriole (Fig. 1.1). During cell division two centrioles appear. They separate so

that each daughter cell receives one. The role of the centriole in cell division is discussed in Chapter 8.

Vacuoles

Many, but not all, cells contain one or more spherical structures in the protoplasm which are termed vacuoles (Fig. 1.1). Vacuoles have a membrane termed the **tonoplast,** and they contain a fluid called the **cell sap.** In plants, vacuoles apparently play a role in determining the color of the flower, fruit, or leaf.

In animals, vacuoles undoubtedly serve to transport substances into and out of the cell. They also may concentrate and otherwise alter the contained substances. Whether this latter function is limited to simple organisms, such as protozoa, or is more general is not known.

Cilia and Flagella

Some cells have hairlike processes capable of vibratory or lashing movements. Numerous short processes are termed cilia (Fig. 1.4).
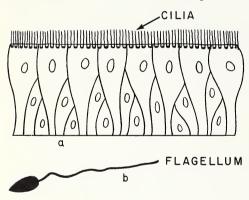
Fig. 1.4. a. Cells Containing Short Hairlike Processes Called Cilia. b. A Cell with a Long Tail-like Structure Called a Flagellum.

If there is but one, long, tail-like structure it is a flagellum. Both cilia and flagella serve to propel the cells through a fluid medium. But there are also ciliated cells which make up stationary tissue. The cilia then serve to clean the surface. For example, the nasal mucosa is lined with ciliated epithelium. It is the cilia which move foreign particles into the pharynx.

Plastids

A plastid is a small body of specialized protoplasm found in many cells, but most notably in plant cells. Various types of plastids are recognized: chloroplasts, leucoplasts, and chromoplasts. The **chloroplasts** are concerned with photosynthesis (Chapter 5). A colorless plastid that plays a role in glucose and starch metabolism is termed

a **leucoplast.** Plastids with red and yellow pigment in them are called **chromoplasts.** Plastids are apparently able to change from one type to another and they can obviously change their number quite independently of cell division. Their composition is so similar to that of mitochondria, it is possible that they are derived from mitochondria. This conclusion, however, is not unanimously accepted, and since there are many notable differences between plastids and mitochondria, there may, indeed, be no relationship between them. Plastids contain enzymes, a fact which suggests that these plastids function in starch, fatty acid, and protein synthesis, as well as in photosynthesis.

CELL MEMBRANE

Plant cells are enclosed by the **cell wall,** animal cells by the **plasma membrane.** These limiting and containing structures are not merely inanimate structures. On the contrary, they play such varied and essential roles in the physiology of the cell that they will be considered in detail in Chapter 3. Here only a general description will be given to complete this outline of the organization of the cell.

Cell Wall

The morphology of the plant cell wall is as varied as are the cells themselves. Some cell walls are thick and rigid, others extremely thin and delicate. In any case it is a supporting and protecting framework. It surrounds the protoplasm and plasma membrane (see below). The cell wall in most plants is considered to be a secretion product of the cellular protoplasm. Three regions, or layers, can usually be identified in the typical cell wall (Fig. 1.5). The outer layer is composed of an intercellular substance which is shared by contiguous cells. It is termed the **middle lamella.** Next is the **primary wall** and inside of that is the **secondary wall** which is generally the thickest layer. It generally contains cellulose and is a tough, protective structure.

Plasma Membrane

In animal cells, the limiting plasma membrane is not supported by the additional cell wall. The plasma membrane is specialized protoplasm. It is generally very thin and difficult to visualize in the unstained cell. The plasma membrane has a degree of elasticity, and

in case of damage it has the power of repair. As described above, the typical plant cell wall has an outer layer, the so-called middle lamella that binds one cell with another. Animal cell membranes do not have this layer but the cell itself secretes a substance that serves the same purpose.
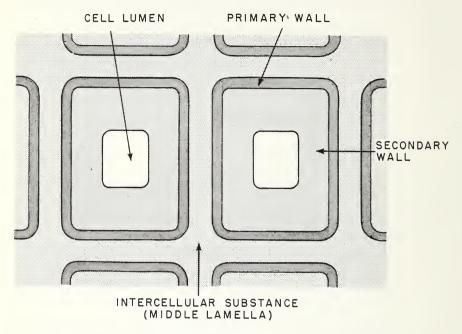
Fig. 1.5. A Diagrammatic Representation of the Cell Wall. Note the three layers: a) the outer layer composed of intercellular substance, b) the middle layer called the primary wall, and c) the inner layer called the secondary wall.

SUMMARY

The basic unit of all living organisms is the cell. The body of the cell is composed of **protoplasm** which is a complex, colloidal, translucent substance. The protoplasm that lies between the nucleus and the cell membrane is often referred to as **cytoplasm.** In the cytoplasm there is a very fine reticular structure termed the **ergastoplasm, or endoplasmic reticulum.** It is thought to be derived from the nuclear membrane. **Microsomes,** considered to be fragments of ergastoplasm, are small granules composed of protein, phospholipid and ribonucleic

acid (RNA). If the cytoplasm is cleared of all its structures by centrifugation, what remains is a clear fluid termed the **ground substance,** or **hyaloplasm.**

Mitochondria, also called **chondriosomes,** may be either rod-shaped or spherical and are enveloped by a membrane. They are composed of protein, fat, phospholipid, and a small amount of RNA. Mitochondria are believed to be essential to the cellular metabolism of oxygen, carbohydrate, fatty acids, and amino acids. They have been considered to be the principal power plant of the cell.

The **Golgi apparatus,** also called Golgi **material** or **complex,** is a relatively large structure composed of empty vacuoles, a network of strands, and small granules all enclosed by a membrane. It is composed of lipoprotein, RNA, and an alkaline phosphatase. It seems to be an intracellular pump that regulates the movement of fluids in the cell and the expulsion of secretory products from the cell.

The **nucleus** is a vital constituent of the cell. It is a large spherical structure composed of specialized protoplasm, termed **karyolymph,** and enclosed by the **nuclear membrane.** Throughout the nucleus there is a network of **chromatin.** During cell division the chromatin network becomes differentiated into **chromosomes.** Deoxyribonucleic acid (DNA) and RNA are constituents of the nucleus. Most of the DNA is in the chromatin while the RNA is in the **nucleolus.** DNA is associated with protein to form **nucleohistone.** There are also present many enzymes, thus indicating the varied roles of the nucleus.

Two types of chromatin are recognized: 1) **euchromatin** and 2) **heterochromatin.** The **genes** are thought to be in association with the euchromatin. Each chromosome is longitudinally subdivided into **chromonemata.** Within the chromosome there are distinct regions, termed **chromomeres,** and a clear, constricted region called the **centromere.** Stained chromosomes have a striped appearance. The dark bands are disks of DNA which may actually be genes.

The nucleus is essential for cell survival. It plays a role in protein and carbohydrate metabolism, and it is of vital importance to cell division and heredity.

Other cell constituents include: a very small granule located close to the nucleus, termed the **centriole;** a spherical structure, the **vacu-**

ole; and **plastids,** small bodies of specialized protoplasm found in many plant cells. Some cells possess **cilia** or **flagella.**

In animals, the limiting cell structure is the **plasma membrane.** Plant cells have an additional **cell wall.**

Problems

1. Make a drawing of a typical cell including the major components.
2. Differentiate between the terms protoplasm, cytoplasm, hyaloplasm and karyolymph.
3. How do mitochondria differ from the Golgi apparatus?
4. What types of nucleic acid are in the nucleus? Where are they most concentrated and what function do they have?
5. Make a drawing showing the structure of a typical chromosome.

Additional Reading

Bourne, G. H., ed., "Cytology and Cell Physiology," Clarendon Press, Oxford, England, 1951.

Brachet, J., "Biochemical Cytology," Academic Press, New York, N. Y., 1957.

Green, D. E. and Hatefi, Y., "The Mitochondrion and Biochemical Machines," *Science,* **133,** 13-19 (1961).

Oberling, C., "The Structure of Cytoplasm," *International Review of Cytology,* **8,** 1-32 (1959).

Roodyn, D. B., "A Study of Metabolic Studies on Isolated Mammalian Nuclei," *International Review of Cytology,* **8,** 279-344 (1959).

Swanson, C. P., "Cytology and Cytogenetics," Prentice-Hall, Inc., Englewood Cliffs, N. J., 1957.

PROTOPLASM

In the previous chapter the general structure of the cell was outlined. Protoplasm, under a variety of names, was everywhere: in the plasma membrane, in the nucleus, and between the nucleus and the membrane. From this observation alone it becomes obvious that the term protoplasm covers a wide variety of substances with a complex and alterable structure, and with myriad functions. It is unarguable that protoplasm is vital to the cell and, therefore, a study of protoplasm is essential to a comprehension of cell physiology.

THE CHEMICAL NATURE OF PROTOPLASM

One approach to an understanding of protoplasm is to clarify its chemical make-up. Thus there has been a systematic attempt to determine the elements and compounds present in protoplasm.

Hydrogen Ion Concentration

Fundamental to a chemical study of protoplasm is the acidity or alkalinity, that is, the intracellular hydrogen ion concentration. Very small changes in hydrogen ion concentration cause important alterations in the living cell. Accordingly, it is essential to be able to determine the hydrogen ion concentration, in the living cell, accurately.

With currently available methods it is relatively easy to determine the acidity or alkalinity of solutions. The main difficulty in ascertaining the intracellular hydrogen ion concentration is that it changes

with alterations in the cell. In other words, one does not simply gather a mass of cells, rupture the cell membrane, collect what comes out and then determine the acidity. The measurement must be carried out in such a way that the cell is altered by the procedure to the least possible extent.

It is common practice to express the hydrogen ion concentration of any substance in terms of pH which is defined as the negative logarithm of the hydrogen ion concentration. This is simply a means of converting unwieldy numbers into a more convenient system. When pH units are used, a neutral solution has a value of 7. Values below pH 7 indicate degree of acidity; above pH 7 degree of alkalinity (Table 2.1).

TABLE 2.1. Hydrogen Ion Concentration Expressed as pH.
$$pH = -Log\ [H+]$$

$[H^+]$	$Log\ [H^+]$	$-Log\ [H^+]$	
Normality		pH	
10^0	0	0	
10^{-1}	-1	1	acid
10^{-2}	-2	2	
10^{-3}	-3	3	
10^{-4}	-4	4	
10^{-5}	-5	5	
10^{-6}	-6	6	
10^{-7}	-7	7	neutral
10^{-8}	-8	8	
10^{-9}	-9	9	
10^{-10}	-10	10	
10^{-11}	-11	11	alkaline
10^{-12}	-12	12	
10^{-13}	-13	13	
10^{-14}	-14	14	

The early attempts to determine protoplasmic pH were carried out in cells that were ground up after freezing. Values of about 5.8 to 6.0 were reported, that is, slightly acid. These efforts were then followed by the insertion of electrodes directly into the normal cell. At first large cells, such as frog or salamander eggs, were used. As microelectrodes, having micron-size tips, were developed, the use of

electrode insertion was extended to smaller cells. Unfortunately, even this procedure is objectionable on the basis that the penetration by the electrode is thought to alter the pH. The studies involving insertion of an electrode into a normal cell gave values that varied all the way from about 6.8 up to 8.5.

Another procedure for determining intracellular pH is to use an indicator. An indicator is a substance which, by its color indicates the pH of the solution to which it is added. Indicators, in very small quantities have been caused to enter the cell. The resulting colors, for different cells, have given pH values ranging from about 5.2 to 6.0.

When the entire problem is surveyed, it becomes clear that no definite statement as to the exact intracellular pH can be made. Quite obviously the pH varies from cell to cell, and from moment to moment in the same cell. In addition, the value obtained depends to a disturbing extent upon the method used. One would only be justified in stating, and then with great caution, that the pH of proto- plasm appears to be close to neutral, i.e., 7.

Protein

One of the main constituents of protoplasm is protein. Protoplasm contains about 15 percent protein.

Proteins, it should be recalled, are complex molecules usually of very great molecular weight. Fundamentally, a protein molecule con- sists of carbon, oxygen, hydrogen, and nitrogen woven together in an intricate pattern. Other elements may also be present in the protein molecules.

Composition. The fundamental building block of protein is amino acid. An example of a simple amino acid is glycine which has the following formula:

$$
\begin{array}{ccc}
& H & O \\
& | & \| \\
H- & C- & C-OH \\
& | & \\
& N & \\
& H_2 & \\
\end{array}
$$

It is the NH_2 group and the COOH ending that characterize amino acids. The general notation for any amino acid is:

$$R-\underset{\underset{H_2}{\overset{|}{N}}}{\overset{\overset{H}{|}}{C}}-\overset{\overset{O}{\|}}{C}-OH$$

The symbol R stands for a carbon chain of variable length.

Amino acids become linked together to form a **peptide.** They unite in the following manner:

$$R-\underset{H-N-H}{\overset{H\quad O}{C-C}}-[OH] \;+\; R-\underset{H-N-H}{\overset{H\quad O}{C-C}}-OH \longrightarrow R-\underset{H-N-H}{\overset{H\quad O}{C-C}}-N-\underset{H\quad H}{\overset{R\quad O}{C-C}}-OH + HOH$$

It can be seen that union is made between the CO group of one amino acid and the NH group of the other. Accordingly, the

$$-\overset{\overset{O}{\|}}{C}-\underset{\underset{H}{|}}{N}-$$

union is termed a **peptide group,** or **bond.**

Peptides form more complex molecules called **peptones** which in turn form **proteoses** and these make up protein. In other words, the amino acid is the basic building block out of which peptides, proteoses, and finally, the finished structure, protein, are built. This is not to imply that the synthesis of proteins is a building process starting with basic amino acids and proceeding through peptides, peptones, and proteoses. To the contrary, proteins seem to be synthesized as a unit rather than in steps. On the other hand, when protein undergoes hydrolysis, the various intermediary products are obtained.

Size. The very great size of the protein molecule is responsible for some important physical properties. Because of this great size, its rate of diffusion is very slow. And for the same reason protein molecules can pass through a membrane only if that membrane has pores of equally large size. Most physiological membranes do not have such

large pores and therefore proteins do not pass through. Accordingly, proteins are primarily responsible for the osmotic attraction of fluid into the cell. In other words, if a cell is placed in pure water, many constituents of the protoplasm will pass through the membrane into the water. But because the protein molecule cannot pass, the water enters the cell. And finally, due to the protein content of protoplasm it exhibits many of the properties of a colloidal solution which are discussed later in this chapter.

Ionic Characteristics. Since amino acids bear both positive and negative charges amino acids are said to be **zwitterions.** The word Zwitter is German and means hermaphrodite, that is an organism with both male and female characteristics. It can be seen that it is not truly an accurate term to describe the negative and positive characteristics of amino acids. The term, **dipolar ion,** has been suggested, but zwitterion persists. Since proteins are composed of amino acids, they have zwitterion characteristics. The relative number of negative and positive charges in a protein molecule determines its **isoelectric pH.** The isoelectric pH is the pH of electrical neutrality, that is, the pH at which the protein will not migrate to the anode or cathode in an electrical field. This isoelectric pH varies from protein to protein.

Proteins undergo ionization (see Chapter 17). If a protein is placed in a solution with a pH more acid than its isoelectric pH, it will ionize as a base. However, in a solution of higher pH (more alkaline), it ionizes as an acid. Accordingly, proteins are said to be **amphoteric** compounds, that is, they ionize both as acids and as bases.

It should also be noted that when a protein ionizes in a medium alkaline to its isoelectric pH the negative charges predominate and therefore it now behaves as an anion. It will migrate to the anode in electrolysis and combine with cations. However, when protein ionizes in a medium acid to its isoelectric pH, the positive charges predominate and it behaves as a cation. Thus with the alteration of the pH of the solution the behavior of the protein is radically altered. In fact, it has been shown that the various properties exhibited by proteins are minimal at the isoelectric pH. As the pH of the solution is changed in either direction, these properties become more pronounced. For example: the osmotic pressure exerted by a protein solution increases progressively as the pH varies in either direction from the isoelectric pH. This is explained on the basis of the fact

that the osmotic pressure of a solution is determined by the number of ions present. Thus, since the degree of ionization increases as the pH of the solution varies from the isoelectric pH, there is a corresponding increase in osmotic pressure.

Other Characteristics. The solubility of proteins is markedly altered by the addition of various salts, especially the salts of the heavy metals. Consequently, proteins may be **salted out** of solution, that is, caused to precipitate by the addition of such inorganic compounds. Another important characteristic of protein is that it is readily coagulated by the application of heat, or the addition of alcohol to the solution. Coagulation simply means the transformation from a fluid state to a compact jellylike mass; however, in the process of coagulation, irreversible chemical alterations occur in the protein molecule. Such alteration is termed **denaturization**, or simply denaturation.

Protein Combinations

Nucleoproteins. As has already been emphasized, the nucleoproteins are extremely important constituents of the cell. They exist and play major roles not only in the protoplasm of the nucleus, but
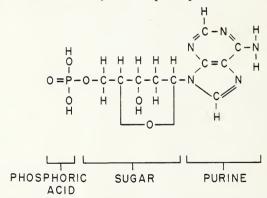

Fig. 2.1.   The Basic Structure of a Nucleotide.

in the protoplasm of the rest of the cell as well. As the term indicates, nucleoproteins consist of nucleic acid in combination with some form of protein.

The basic building unit of nucleic acid is the **nucleotide**. It can be seen in Fig. 2.1 that a nucleotide consists of a sugar and phosphoric acid in combination either with a pyrimidine or purine base. The

two types of nucleic acid of importance in cell physiology, as already mentioned, are: ribonucleic acid (RNA) and deoxyribonucleic acid (DNA). The composition of each is shown in Table 2.2.

TABLE 2.2. Nucleic Acid Composition

RNA	DNA
Phosphoric acid	Phosphoric acid
Sugar: ribose	Sugar: deoxyribose
Purine: adenine guanine	Purine: adenine guanine
Pyrimidine: cytosine uracil	Pyrimidine: cytosine thymine

Just as amino acid is the structural unit of protein, the nucleotides are the building units of nucleic acids. For this reason the larger nucleic acids are sometimes referred to as polynucleotides. The linkage is thought to be between the phosphate residue of one nucleotide and the sugar component of another. Exactly how nucleic acids link with protein to form nucleoprotein is not clearly understood. Whatever the exact nature of the linkage it is apparently a weak one, since it may be readily broken.

Other Protein Combinations. Proteins enter into interesting unions with substances other than the nucleic acids. For example there are many **lipoproteins**, that is, structures that are a combination of protein with lipid. Lipoproteins seem to play a major role in cell structure. Further, proteins unite with glycogen and with various anions and cations. The physiological significance of these combinations is great and will be discussed in subsequent chapters.

Carbohydrates

In protoplasm, carbohydrates exist both in the free state and in combination with protein in the form of nucleoprotein. Carbohydrate is one of the basic substances of protoplasm and in various forms enters into vital cellular functions.

Carbohydrates, it will be recalled, are composed of carbon, hydrogen, and oxygen. The basic pattern is CH_2O, that is, two hydrogen atoms to one carbon and one oxygen. The simplest form of carbohydrate is the **monosaccharide**, which may contain from 2 to 10 carbon atoms. The most common monosaccharide is glucose ($C_6H_{12}O_6$). **Disaccharides**, as the term indicates, are formed by the combination

Fig. 2.2.   Sucrose.

of two monosaccharides. Since in this reaction there is the loss of a molecule of water (Fig. 2.2), the basic carbohydrate pattern of CH_2O no longer holds. Sucrose, for example, has the formula $C_{12}H_{22}O_{11}$. The complex sugars, termed **polysaccharides**, result from the combination of more than two monosaccharides. Glycogen and cellulose

Fig. 2.3.   Ribose and Deoxyribose.

are two common examples. Finally, it should be recalled that a monosaccharide with five carbon atoms is called a **pentose**; one with six carbon atoms is a **hexose**. The pentose sugars, ribose and deoxyribose, are of great physiological importance in the form of RNA and DNA. Their structural formulas are indicated in Fig. 2.3. It can be

noted that the only difference between RNA and DNA is that the deoxy form is minus an oxygen atom. As the loss generally takes place at the second carbon atom, the correct designation is 2-deoxy-ribose.

The role of carbohydrates in cell physiology will be discussed in detail later. Here it is sufficient to mention that the simple sugars constitute an immediate source of food, that is, of energy for the cell, whereas the polysaccharides represent reserve sources. They can be utilized for energy, but first they must be broken down into simpler forms.

Lipids

The term lipid includes neutral fats, fatty acids, fatty oils, phosphatides, waxes, and steroids.
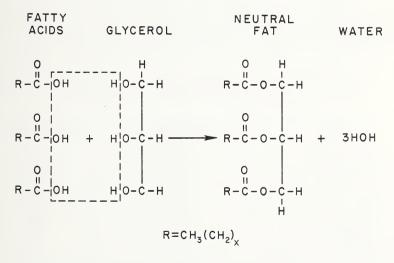

$$R = CH_3(CH_2)_x$$

Fig. 2.4.   Combination of Fatty Acids and Glycerol to Form Neutral Fat.

The neutral fats constitute the most common type of lipid. They are composed of fatty acids and glycerol. The basic combination is three molecules of fatty acid with one of glycerol (Fig. 2.4). The next largest group of lipids is made up of phosphatides, also called **phospholipids.** Unlike neutral fats, the phosphatides are soluble in fat solvents and also in water.

The steroids have diverse physiological activities that are of special significance. They are all basically composed of a fused ring:

The steroids differ in having various side chains. Among the more important steroids are cholesterol, ergosterol, and a large group of hormones.

Cations and Anions

The concentration of potassium, sodium, calcium, magnesium, chloride, bicarbonate, sulfate, and phosphate have been determined in various types of cells and body fluids. It is remarkable that there is

TABLE 2.3. Intracellular and Extracellular Fluid Composition.*
(mEq/liter H_2O)

Ion	Intracellular	Extracellular
Potassium	120	5
Sodium	5	145
Calcium	?	4
Magnesium	30	3
Chloride	5	105
Bicarbonate	10	28
Phosphate	98	3
Sulphate	20	1

* Average values which differ, sometimes widely, in different cells.

a recurring pattern. The concentrations of these ions within the cell do not vary greatly for most cells, but are quite different from the concentrations of these ions in the fluid surrounding the cell (Table 2.3). In general, it is found that potassium is much more concentrated within the cell than without while sodium is just the opposite. There is usually more magnesium in the protoplasm than in the surrounding fluid. As for the anions, phosphate appears as the main proto-

plasmic constituent while chloride is much more concentrated extra-cellularly. Alteration in the concentration of these cations and anions usually results in grave physiological consequences.

In addition to the ions already listed, others are often found in small quantities in protoplasm. They include iron, copper and manganese, and in even smaller amounts, zinc, nickel, lithium, rubidium, iron, molybdenum, and vanadium.

Water

About 80 percent of protoplasm is water. The exact percentage varies in different types of cells.

Cellular water is said to be **bound,** that is, it is in chemical combination with other constituents of protoplasm, primarily protein. The question is, what happens when dried or crystalline proteins are placed in solution? Is the water still bound? Some authorities contend that in the normal protoplasmic state bound water does not exist. That is to say, it is **free** and therefore available for the countless number of cellular functions. This question of the existence of bound water has not yet been resolved. The consensus is growing that most of the cell water is not bound.

THE PHYSICAL NATURE OF PROTOPLASM

Knowledge of the chemical make-up of protoplasm contributes very little to a comprehension of what it is about the substance that permits it to be the basis of life, so to speak. It is quite simple to make a mixture of protein, carbohydrate, lipid, water, and the various ions, yet it does not closely resemble protoplasm nor does it come alive. Every student of cell physiology is ultimately faced with the question: what is life? That there is no direct, simple answer to this question is made obvious by asking still another question: are viruses alive? If the answer is yes, then it is interesting to note that some viruses have been reconstituted from chemical components. It may further be noted that many of the characteristics of life, such as genetic determination of cell morphology and biochemistry, have been duplicated by the use of chemical substances, such as DNA. The point being made is that life is as complex and as difficult to define as is protoplasm. There appears to be something in the bio-

logical system which transcends the mere mixture of biological molecules. The properties we attribute to life may be due in part to a complex stabilization of a particular physical state. For this reason it is essential to investigate the physical nature of the intracellular contents.

Viscosity

The term viscosity may be defined as a manifestation of molecular attraction. It is the property of fluidity. A highly viscous substance is one that offers great resistance to a change in form. This property of viscosity is derived from the molecular attraction between the molecules of the substance. The more viscous a substance, the less readily it flows. Protoplasm may be considered to be highly viscous, at least in comparison with water.

Viscosity can be measured by determining the rate of flow of the fluid through a tube. The higher its viscosity, the less readily will it flow through a tube. Thus the relative viscosity of a fluid, relative usually to water, can be easily determined. Obviously, this method cannot be used to ascertain the viscosity of protoplasm, since the very act of forcing protoplasm through a tube drastically alters it.

A method for measuring viscosity that has been widely used is based upon Stokes' law. It is interesting to note that Stokes has been honored for his work in the field of viscosity by having a new unit of viscosity named for him, the **stokes.** The stokes equals the viscosity of a substance expressed in poises divided by its specific gravity. The **poise** is defined as the absolute viscosity of a fluid that would require a shearing force of one dyne to move a square-centimeter area of either of two parallel layers of the fluid, one centimeter apart, with a velocity of one centimeter per second, relative to the other layer, the space between the layers being filled with the fluid. A fluid of one stokes, then, has a viscosity of one poise and a density of one gram per cubic centimeter. It is easy to understand why viscosity is usually expressed on a relative basis!

Stokes' law expresses the velocity of a sphere falling through a liquid. This velocity, if all other factors are constant, varies with the viscosity of the fluid. The movement of a small sphere in a viscous substance like protoplasm is very slow. To hasten it the force moving the sphere can be increased by centrifugation.

Other methods have also been developed. By the use of these various procedures the viscosity of protoplasm in many types of cells has been measured. Values for different cells vary, of course, but in general they seem to be in the range of 2 to 20 centipoises. This is to be compared with water which has a value of about 1 centipoise, that is, one-hundredth of a poise.

Viscosity measurements have revealed significant alterations which are associated with various cellular activities. For example, correlations have been found between changes in viscosity and fertilization of the ovum. Later, when ameboid movement is considered, it will be seen that there are sol-gel transformations, that is viscosity changes, which account for the ability of an ameba to move.

Inherent Structural Bonds

In the previous chapter it was pointed out that protoplasm has a reticular appearance. Although the nature, or structure, of this endoplasmic reticulum is not known, the fact that protoplasm is not simply a fluid is shown in various interesting experiments. For example, a metal sphere may be placed in certain areas of a cell and then attracted by a magnet located outside the cell. The sphere moves due to this attraction, but when the magnet is removed, the sphere is seen to spring back to the position from which it was attracted. It is as though the protoplasm consisted of a fine elastic mesh that is displaced and stretched by the movement of the metal sphere. If the sphere is placed in areas of the cell where the protoplasm appears to be a homogeneous liquid, following displacement by magnetic attraction, the sphere does not rebound. Such experiments indicate that there is a structure with elastic properties only in some areas, usually the cortical layers of the protoplasm.

That there are probably structural bonds of some nature in protoplasm is also shown in another way. In Fig. 2.5 the relationship between applied force and rate of flow is shown. It should be noted that a homogeneous fluid such as water, or even one with high viscosity such as oil, exhibits a straight line relationship between force and flow. In addition it is to be observed that the curves for these types of fluid pass through the zero point for both values. That is to say, as soon as pressure is applied, flow begins. A very different situation exists insofar as protoplasm is concerned. In the first place, it

is obvious that considerable force must be applied before any flow occurs. Second, it can be seen that the line is not straight, at least at low flows. This difference in flow performance has been explained on the basis of structural bonds in protoplasm which oppose flow so that it only occurs when sufficient force is applied to break the bonds. The more force the greater the rupture until ultimately the flow characteristics are identical to that for homogeneous, viscous fluids.
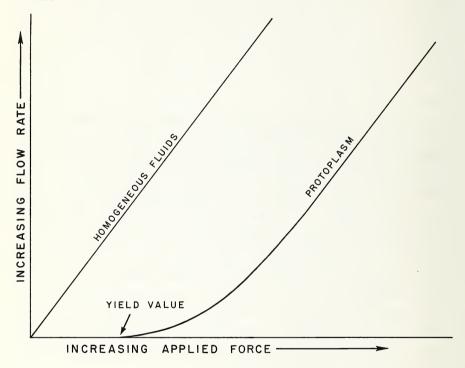

Fig. 2.5. The Relationship of Applied Force to Rate of Flow. Note that considerable pressure must be applied before protoplasm begins to flow.

Light Refraction

Rays of light move through air at a velocity of about 300,000 kilometers per second. Through other media, such as glass or protoplasm, the velocity is less. When light rays pass at an angle through the interface between two media in which the velocity is different, the rays are bent. Such bending is termed refraction. The explanation

lies in the fact that the rays on one side of the interface will travel at a different velocity than the rays on the other side. Accordingly, the beam is bent (Fig. 2.6).

The ratio between the velocity of light in a substance compared with that in air, or a vacuum, is termed the **refractive index.** Thus, the refractive index for air is 1.0, for protoplasm it is about 1.4. Such measurements, it should be understood, assume that protoplasm is homogeneous, which it is not. Certain parts of the cell exhibit what is known as **double refraction,** or **birefrigence.** This means that a ray of light is split into two rays, one of which is refracted and the other is not. A substance that causes double refraction is called **anisotropic.** A substance through which all rays of light pass at the same velocity, and thus does not cause double refraction, is termed **isotropic.**

Fig. 2.6. The Refraction of Light.

Protoplasm itself does not usually exhibit double refraction, but various structures located in the protoplasm often do. These include the nuclear membranes, the chromosomes, an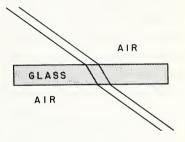d the mitochondria. There are reports of protoplasm, which is normally isotropic, being anisotropic. However, in all such cases the prior treatment of the cell suggests that the protoplasm may have been altered from its normal state.

Much work has been done in this field, and double refraction is now even subdivided into positive and negative categories. Thus, if the velocity of the bent ray is greater than the other ray, one speaks of positive birefrigence; if it is slower, negative birefrigence. It has further been noted that substances containing nucleoproteins generally cause positive birefrigence, whereas lipid substances have the opposite effect.

The significance of investigating the refraction of light as it passes through various cellular structures is that it provides valuable information as to the orientation of their constituent molecules. For example, such studies have contributed much to an understanding of the arrangement of molecules in the cell membrane. As will be discussed later, from these and other investigations the cell membrane

is visualized as consisting of a double layer of fat molecules, covered with a thin layer of protein. Further, the orientation of each molecule is suggested by the manner in which light is refracted.

THE COLLOIDAL NATURE OF PROTOPLASM

Protoplasm is not a true solution. It has long been considered to be, or at least to exhibit, many of the properties of colloids.

Definition

It is difficult to formulate an all-inclusive, yet specific, definition for a colloid. A colloid is usually defined as a state of subdivision of matter in which the individual particles are of submicroscopic size, but yet are larger than ordinary molecules. In other words, such particles although larger than ordinary molecules cannot be seen with the usual light microscope. According to this concept, colloidal particles have an arbitrary size between 1 and 100 millimicrons. But, as noted, this range is purely arbitrary, and quite often solutions with particles of a considerably larger size are considered to be colloids.

Colloidal Properties

Because the particles of a colloid are larger than the molecules and ions of a true solution, they diffuse through the solution far more slowly. In other words, the rate of diffusion of a particle is inversely proportional to its size. Another consequence of colloidal size is that colloidal particles cannot pass through the pores of most membranes.

Due to the fact that the colloidal particles have a different refractive index than the medium in which they are dispersed, rays of light in passing through a colloidal solution are scattered so as to produce a cone of light which appears to have a faint blue cast. This is termed the **Tyndall cone,** and it is characteristic of colloids.

Many of the properties of colloids stem from surface phenomena. Colloidal particles are small, but nonetheless there is a very definite surface to each particle; when the area of this surface is multiplied by the number of particles present in a colloidal solution, a very large number results, indicating a huge surface area. The boundary between the particle and the dispersing medium is termed the **interface.**

These surfaces have an electrical charge and thus there is attraction or repulsion. The forces which result from these huge areas are often remarkable and are directly responsible for colloidal behavior.

Sols and Gels

A sol is a colloidal solution. A sol consists of a dispersion medium and the colloidal substance. A gel is also a colloidal system in which there is a solid and a liquid phase so that the whole mass exists as a solid or a semisolid. According to this somewhat obsolete concept, in a sol the fluid phase is continuous and the particles are dispersed throughout it. A gel, then, has the solid phase continuous and the fluid broken up into discontinuous areas or pockets. So many exceptions exist to this concept of phases that it is hardly used at the present time. Actually, there are all degrees of gradation between gels and sols. According to newer concepts of colloids, sol-gel transformations are thought of in terms of the formation of enough cross bonds between macromolecules so as to give structure to what was previously a liquid.

Protoplasm may exist as a sol or a gel, or even in a single cell both may be present. In many cells the protoplasm in the center has all the properties of a sol. This is then surrounded by a stiff gel-like outer layer, termed the cortex. The cellular membrane is often indistinguishable from the cortical gel and is itself a gel.

The interesting aspect of the cortical gel is that its rigidity apparently depends upon the presence of calcium ions. The viscosity of the cortical gel has been shown to be directly proportional to the quantity of calcium present. Such alterations may well influence intracellular activities and thus supply a clue to the influence of various ions on cellular function.

The existence of protoplasm as a gel is not restricted to the cortex or the plasma membrane. Various structures within the cell are considered to be gels. The centrosomes and chromosomes may be so composed.

The ability of protoplasm to change the colloidal state in which it exists is shown very dramatically by the so-called **surface precipitation reaction.** If a cell membrane is ruptured, some of the protoplasm escapes, but then a new cell membrane quickly forms to prevent the

complete loss and spread of the fluid protoplasm. Again it has been shown that in order for this surface precipitation reaction to occur, calcium must be present. Thus it appears that in some manner, with the aid of calcium, fluid protoplasm is capable of being converted into a gel which, in effect, functions as a new limiting membrane.

The Role of Protein

Most, if not all, of the colloidal properties exhibited by protoplasm are thought to be dependent upon the protein present. Protein molecules are large and are probably dispersed throughout the protoplasmic ground substance to form a colloid. In other words, the ground substance may be thought of as the dispersing medium and the protein molecules as the dispersed particles. On this basis one can account in protoplasm for the Tyndall cone, high viscosity, structural bonds, surface phenomena, and gel formation.

SUMMARY

Intracellular pH has been variously estimated all the way from 5.8 to 8.5. In the living state the pH is probably very close to neutral, i.e., 7. Protoplasm is composed of about 80 percent **water,** 15 percent **protein,** 3 percent **lipid,** 1 percent **carbohydrate,** and 1 percent **salt.**

Proteins are made up of **amino acids** that unite to form **peptides,** then **peptones** and finally **proteoses** before the very large protein molecule is achieved. Due to the size of the protein molecules, diffusion is slow and most membranes will not permit them to pass. Amino acids are termed **zwitterions,** or **dipolar ions,** because they have both negative and positive charges. The relationship of these charges in the protein molecule determines its **isoelectric pH.** Since proteins ionize both as acids and as bases, they are termed **amphoteric.** Proteins are caused to precipitate, to be **salted out** of solution, by the addition of salts of the heavy metals. When protein is caused to coagulate, irreversible chemical alterations occur; this process is termed **denaturization.** Proteins combine with nucleic acid to form **nucleoproteins;** they also unite with lipids and various ions.

Carbohydrates are composed of carbon, hydrogen, and oxygen. The simplest carbohydrates are the **monosaccharides** which combine

to form **disaccharides** and **polysaccharides.** The pentose monosaccharides are of particular importance to protoplasm because they appear in ribonucleic acid and deoxyribonucleic acid.

Lipids include **neutral fats, phosphatides, waxes,** steroids, and other substances. Neutral fats are composed of three molecules of **fatty acid** and one of **glycerol.** The more important steroids are **cholesterol, ergosterol,** and many **hormones.**

Protoplasm is viscous. The units of viscosity are the **poise** and the **stokes.** More often, viscosity is expressed simply in relation to water. The viscosity of protoplasm varies from cell to cell in the range of about 2 to 20 centipoises, that is, 2 to 20 times the viscosity of water.

Protoplasm has very definite structural bonds that resist the movement of objects through it. These bonds also resist flow so that considerable pressure must be applied to protoplasm before flow begins.

Light moves more slowly in protoplasm than it does in air, thus rays of light are refracted. Some parts of protoplasm split rays of light into two beams, only one of which is bent. This phenomenon is termed **double refraction,** or **birefrigence.**

Many of the properties of protoplasm are best understood by considering it to be a **colloid.** The colloidal properties of large particles, the **Tyndall cone,** and the surface phenomena are probably mostly or wholly attributable to the protein present in protoplasm.

When the cell membrane is ruptured, the complete loss of protoplasm is prevented by the rapid **surface precipitation reaction** which depends upon the presence of calcium.

Problems

1. Outline the difficulties in determining the pH of protoplasm.
2. Since only about 15 percent of protoplasm is protein, explain why this constituent is considered so important.
3. How does ribonucleic acid differ from deoxyribonucleic acid?
4. How may the viscosity of protoplasm be determined?
5. Define the following terms:

   a. Polysaccharide   d. Amphoteric
   b. Birefrigence   e. Anisotropic
   c. Surface precipitation reaction   f. Peptide

Additional Reading

Davidson, J. N., "The Biochemistry of the Nucleic Acids," 3rd ed., John Wiley & Sons, Inc., New York, N. Y., 1957.

Lovern, J. A., "The Chemistry of Lipids of Biochemical Significance," Methuen & Co., Ltd., London, England, 1955.

Neuberger, A., ed., "Symposium on Protein Structure," John Wiley & Sons, Inc., New York, N. Y., 1958.

Pigman, W. W., ed., "The Carbohydrates: Chemistry, Biochemistry and Physiology," Academic Press, Inc., New York, N. Y., 1957.

Scheer, B. T., "General Physiology," John Wiley & Sons, Inc., New York, N. Y., 1953.

THE CELL MEMBRANE

IN VISUALIZING the cell membrane, one pictures a structure that limits and contains the protoplasm—a protective and supporting structure somewhat akin to the walls of a room, or better the walls of a honey-comb. This general picture is, unfortunately, not always the case, and, as a matter of fact, in many animal cells it is far from the truth. One speaks of the cell membrane as though it were a definite and distinct entity, but in many cases it is not. In some cells one can, to be sure, see the cell membrane, separate it from the rest of the cell, measure it, analyze it, and in every way consider it and treat it as a separate structure. But there are also cells in which the limiting area of the cell is indistinguishable from the rest of the protoplasm. Thus, although in this chapter the functional anatomy of the cell membrane is considered as though it were a separate and distinct cell structure, the above mentioned variations should be kept clearly in mind. Because of this variation, some authorities prefer to speak not of a cell membrane, but merely of the protoplasmic surface.

ANATOMY OF THE CELL MEMBRANE

In Chapter 1 it was pointed out that the limiting structure in plant cells is called the cell wall, whereas in animal cells it is spoken of as the plasma membrane. It was also indicated that some cells are sur-rounded by additional layers, sometimes termed coats, or extraneous membranes. These extraneous layers generally serve a protective function and can be removed by various means without altering the cell, or the underlying cell membrane. Such protective extraneous
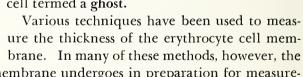
layers permit substances to leave and to enter the cell. This is, on their part, a passive process. It is the true cell membrane that plays an active, dynamic role in such transfer.

Cell Membrane Thickness

Since the limiting protoplasmic surface is often indistinguishable from the rest of the protoplasm, it is clearly impossible to measure the thickness of the cell membrane of such cells. In those cells one simply speaks of a plasma membrane in a functional and not an anatomical sense. There are, however, cells which do have a very definite membrane.

The favorite cell for membrane studies is the **erythrocyte,** that is, the red blood cell. As there are some 5,000,000 erythrocytes per cu mm of blood in most mammals, there is an abundant and convenient supply. They are also quite large. Human erythrocytes have a diameter of about 7.2 microns and a thickness of approximately 2.2 microns near the circumference, and about 1 micron at the center (Fig. 3.1). Before a study of the membrane of erythrocytes is made, the cells are placed in hypotonic solutions, and, as a result, they lose their hemoglobin which occupies much of the cell. What remains after such treatment is an almost empty cell termed a **ghost.**

Fig. 3.1. The Dimensions of an Erythrocyte.

Various techniques have been used to measure the thickness of the erythrocyte cell membrane. In many of these methods, however, the treatment that the membrane undergoes in preparation for measurement may drastically alter its thickness. It is for this reason that such a wide range of results has been reported. One finds values in the literature ranging from about 50 Ångstrom units all the way up to 5000 Ångstroms. And these values are for the same type of cell! The latter figure was reported for the undried membrane and may be very close to the actual thickness of the cell membrane in the living state. It is known that the thickness of the erythrocyte across the center is about 1 micron. If the two membranes were in contact at
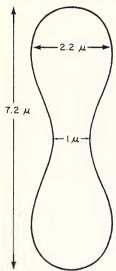
this point, then each one would have to be 0.5 micron, that is, 5000 Å thick. But this high value for the erythrocyte cell membrane has been doubted by many authorities. The thickness most often reported using a variety of methods is about 150 Å. One conclusion is clear: if so little agreement exists as to the thickness of the relatively large erythrocyte cell membrane, there is little possibility of agreement as to the cell membrane thickness in smaller cells in which the membrane is far less distinct.

Chemical Composition

Erythrocyte ghosts are the favorite subjects for the analysis of the chemical composition of the cell membrane. The cell membranes of bacteria have also been studied. There is general agreement that protein and lipid are the major constituents; however, calcium is also present and apparently is essential to the function of the membrane. It will be recalled that in the absence of calcium the surface precipitation reaction does not occur. Furthermore, it has been shown that the calcium content of the membrane to some extent determines its permeability to water.

The protein of the cell membrane has a characteristic nature which has been termed **stromatin;** it is fibrous and has a high molecular weight. Analysis of the surface precipitation reaction suggests that the cell-membrane protein forms by differentiation from cytoplasmic protein.

There is usually somewhat more lipid in the erythrocyte cell membrane than protein, although they may appear in equal concentration. The lipid consists of cholesterol and phospholipids, mainly lecithin and cephalin.

In addition to lipid, protein, and calcium, water and other ions are also present in the membrane.

Cell Membrane Structure

There is general agreement as to the composition of the cell membrane, but just how these components are arranged to form that membrane is a vastly different problem. It should be understood that even with the electron microscope, one cannot specifically discern the membrane structure. Thus, even though rather elaborate and impressive pictures purporting to be the structure of the cell membrane

have been drawn, knowledge on the subject is nowhere near that clear. In short, the structure of the membrane has been deduced from various lines of evidence.

The surface tension, the resistance, and the passage of various substances through the membrane have been extensively studied. These investigations plus those involving X-ray diffraction have given rise to the belief that the cell membrane, at least the erythrocyte membrane, has a structure consisting of two layers of fat molecules (Fig. 3.2). The arrangement of the protein is a greater problem. It has
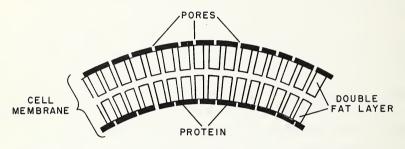
Fig. 3.2. Theoretical Cell Membrane Structure. Showing a double layer of fat covered by a thin layer of protein.

been suggested that because of the great surface activity of protein that any lipid surface in contact with a protein solution will be covered by the protein. Thus, it is thought that the double layer of fat molecules is covered on both sides by a very thin layer of protein.
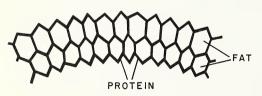
Fig. 3.3. Theoretical Cell Membrane Structure. The protein is thought to form a honeycomb which contains the fat.

The question of the pores in the membrane remains unsolved. There seems to be little doubt that pores do exist, but whether they appear between molecules of protein, or molecules of fat, or between both is uncertain. Because the passage of various substances through membranes seems to depend on at least two factors: 1) molecular size, and 2) solubility of the substance in fat, there is the assumption that the pore size varies and is determined, on the one hand, by the distance between the protein molecules, and, on the other, by the dis-

tance between the fat molecules. Thus, a relatively large molecule could penetrate if it could fit into the opening between the protein molecules and then be soluble in the lipid. On the other hand, if the molecule is not soluble in lipid then it must be small enough to pass through the smaller openings between the lipid molecules.

The above concept is convenient but more direct evidence indicates that the membrane may be built like a honeycomb with the protein forming the structure and fat filling the openings (Fig. 3.3).

PHYSICAL PROPERTIES OF THE CELL MEMBRANE

Surface Charge

That the surface of the living cell has an electrical charge can be easily demonstrated. If the cells are suspended in a solution through which a current is passed, the cells may move to the anode. Such movement, of course, indicates that the cell surface has a negative charge. There are some cells, such as protozoa and spirochetes, that migrate to the cathode, indicating that they have a positive charge.

The movement of cells in response to an electric current is thought to be due, practically exclusively, to the charge on the surface of the cell and not to any influence of the internal cytoplasm. It is probably the membrane protein that carries the surface charge.

Surface Tension

In general, the term tension involves the concept of resistance to being stretched. There is an inter-molecular attraction which is termed cohesion. Because of these cohesive forces surface molecules oppose being separated one from the other. This opposition, this resistance, is known as surface tension.

In a cell, similar forces operate to hold it together, to maintain a spherical shape, to oppose distortion. Thus, even if no cell membrane existed, the cohesive molecular forces would still exert a degree of surface tension. When a definite cell membrane is present, the surface tension is greatly increased.

Surface tension is expressed in dynes per centimeter. Water in air has a value of about 73 dynes per centimeter. A drop of oil covered with a thin protein layer and surrounded by water has a surface tension of only about 1 dyne per centimeter. Thus, it is not surpris-

ing to find that many cells which have a lipid-protein membrane and exist in a watery medium have a surface tension that is very low, that is, usually 1 dyne per centimeter or less.

MOVEMENT OF WATER THROUGH THE CELL MEMBRANE

The cell membrane provides physical protection for the cell, but it has an equally important function in that it controls the movement of all the substances that the cell requires, as well as the waste products resulting from intracellular metabolic processes. Water is one such vital substance, the movement of which must be rigorously regulated for cell survival. The movement of water through the cell membrane depends upon many forces in relation to the membrane characteristics. These forces include hydrostatic and osmotic pressure.

Molecular Activity

Molecules of water at all temperatures above absolute zero are in constant movement. Thus, if a membrane containing pores large enough for the water molecules to pass through is placed in water, there will be a movement through the membrane in both directions simply by diffusion. This is a random process which depends upon the chance encounter of a molecule with a pore. Since the molecular movement is proportional to the temperature, the rate of movement through a membrane due to diffusion is also proportional to the temperature.

Hydrostatic Pressure

When water is poured into a funnel in which filter paper has been fitted, the water quickly moves through the filter paper. The force driving the molecules through the pores of the paper is due to gravity. Likewise, if a membrane separates two compartments of differing hydrostatic pressure, there will be a movement of water from the high pressure compartment to the low pressure compartment. This type of transfer is characteristic of the movement of blood plasma through the capillary wall into the tissue spaces. Since the hydrostatic pressure in the capillary is usually higher than it is in the tissue spaces, plasma is forced through the wall.

Osmotic Pressure

Cells behave like osmometers. The membrane is permeable to some molecules but not to others. In addition, protoplasm contains protein which does not readily move through the membrane. Thus, if a cell is placed in a hypotonic solution, it will swell because of the inward movement of water. That is, the greater molecular concentration in the protoplasm attracts water into the cell and therefore the cell swells. Conversely, if the cell is placed in a hypertonic solution, that is, one containing molecules which do not readily pass through the membrane, then water from the protoplasm will pass through the membrane and the cell will shrink.

It should be understood that the movement of water due to osmotic pressure depends upon the permeability characteristics of the membrane. And, at this point, a clear differentiation between a **semipermeable** and a se-

HYPOTONIC →|← HYPERTONIC

HYPOTONIC →|← HYPERTONIC

H₂O   10 PERCENT GLUCOSE   20 PERCENT GLUCOSE
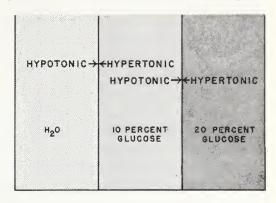
Fig. 3.4. The Relationship of Hypotonic and Hypertonic Solutions.

lectively permeable membrane must be made. A semipermeable membrane permits movement of all molecules which are smaller than the pores in the membrane. A selectively permeable membrane, on the other hand, permits some substances to pass more readily than others, but size is but one consideration which may be outweighed by others. For example, such a membrane may permit a larger molecule to pass more readily than a smaller one. It is thought that all living cell members are selectively permeable.

The terms hypotonic and hypertonic have been used and should be defined. These are relative terms. Thus, if one solution has exactly the same osmotic pressure as another, it is said to be **isotonic** to that reference solution. **Hypotonic** means, literally, below normal tension or strength. In this connotation it describes a solution with lower osmotic pressure than a reference solution. A **hypertonic** solu-

tion has a higher osmotic pressure than the reference solution. For example, as shown in Fig. 3.4, three solutions are separated by membranes. The first compartment contains pure water, the second 10 percent glucose solution, and the third 20 percent glucose solution. The 10 percent glucose solution in the middle compartment is hypertonic with reference to the water, but it is hypotonic with reference to the 20 percent glucose solution.

Experiments in which various types of cells have been placed in hypo- and hypertonic solutions indicate that the cell membrane is generally highly permeable to water. In fact, water enters so freely that in very hypotonic solutions the inward movement continues until rupture of the cell occurs.

Permeability Units

In order to compare the movement of water through various cell membranes it is necessary to have a unit to express permeability. Various such units have been suggested, but the unit, or constant, that has been most widely used expresses permeability in terms of the number of cubic microns of water entering the cell per minute per unit area of membrane per atmosphere of difference of osmotic pressure between the internal and external media. This rather involved sounding definition simply means that the amount of water that enters the cell depends upon the surface area of the cell and the difference in osmotic pressure between the cell protoplasm and the surrounding medium. In these units it has been found that the permeability of various cells to water varies all the way from about 0.2 to 2.8 cubic microns per minute. The latter figure was obtained using erythrocytes.

MOVEMENT OF SOLUTES THROUGH THE CELL MEMBRANE

The relationship between the size of the solute molecule and the diameter of the pore is only one factor in determining the degree of permeability of that substance. Another important consideration is the degree of solubility of the solute in the substance of the membrane. The forces which influence movement of solutes through membranes are: 1) the concentration gradient and 2) in the case of charged particles, electrostatic considerations.

Partition Coefficient and Molecular Size

The ratio of the solubility of a substance in a lipid solvent to its solubility in water is termed the **partition coefficient.** To put it an- other way, if a solute dissolves readily in lipid but not in water, then it will have a high partition coefficient. The movement of various substances through the cell membrane has been determined and the rate of movement plotted against the partition coefficient (Fig. 3.5).
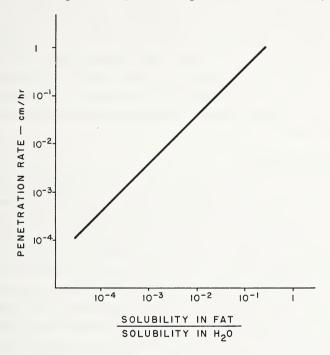

Fig. 3.5. The Relationship of the Partition Coefficient (solu- bility in fat/solubility in water) and Penetration Rate. It can be seen that the more soluble a substance is in fat the more readily it penetrates the cell.

It will be noted that there is a high positive correlation between these two values, which indicates that the more soluble the solute is in lipid the more readily it enters the cell. Rather extensive studies have been carried out with the alcohol series. In these series, the higher alcohols have large molecular weights, but much greater solu- bility in lipid. It is found that, despite the larger molecular weight,

the rate of movement into the cell is greater for the higher members of the series. This is not to say that molecular size is of no importance. If a molecule is small enough to pass through the pores in the membrane, then it enters far faster than would be anticipated from a consideration of its partition coefficient. On the other hand, with large molecules the main determinant appears to be the solubility in lipid. In short, it may be generalized that the deciding factor for the movement of molecules below a critical size is molecular size, but above this size solubility in lipid is the crucial consideration. As is to be expected, there is great variation in this generalization when different cells are studied. Thus, with one particular cell molecular size may prove to be of greater importance than lipid solubility, and the reverse be true for other cells. But in any consideration of the movement of solutes through membranes both factors must be taken into consideration.

Concentration Gradient

If a molecule is capable of free movement through a membrane, the rate at which it will move is determined, if all other factors are constant, by the difference in concentration of that substance on the two sides of the membrane, that is, the concentration gradient. This follows from a simple consideration of diffusion. Again it will be recalled that molecules, at all temperatures above absolute zero, are in constant movement. The more rapidly they move the greater is the chance that they will enter a pore of the membrane and move through. But if the temperature of the solutions on both sides of the membrane is the same, then temperature is no longer a factor. However, if the concentration on one side is greater than on the other, there will be more molecules of the solute in solution on that side. Accordingly more molecules will pass from the concentrated solution to the less concentrated solution than vice versa.

Movement of Ions

The movement of ions through the cell membrane is determined not only by the size of the ion, its solubility in lipid, and concentration gradient, but by still another factor, namely charge. A wide variety of experiments has led to the conclusion that the stronger the

charge, the slower the penetration of the membrane. This means that weak electrolytes, i.e., ones that do not readily ionize, enter more readily than do stronger ones. It also follows that monovalent ions penetrate the membrane more rapidly than do divalent or trivalent ions.

Exactly why the presence of a charge on the ion should make it more difficult for that ion to move through the membrane is not certain. It will be recalled that proteins are zwitterions, that is, they bear both positive and negative charges. Phospholipids have the same characteristic. Accordingly, the cell membrane which is composed of protein and phospholipid represents a barrier in which both charges are present. Thus, as a charged ion moves through the membrane, there will be attraction to the part of the protein or phospholipid bearing the opposite charge, and this is thought to hinder its movement. On the other hand, there are also similar charges present which should result in repulsion. Just why attraction and repulsion do not offset one another and permit ion movement unrestricted by consideration of charge is not clear. It may be the relative position of the charges that is of controlling importance.

The movement of sodium and potassium ions has been studied perhaps more extensively than any other ionic movement. It is found that potassium moves through the cell membrane more rapidly than does sodium. Yet they both have the same charge. The explanation may lie in the fact that in the hydrated form the diameter of the potassium ion is considerably smaller than the diameter of sodium.

The entire problem of the movement of ions through membranes into cells is complicated by the fact that in the normal living state there is a striking imbalance between the concentration of various ions in the protoplasm and in the external environment. It is obvious that other factors besides ionic size and concentration must operate to maintain the high internal potassium and the low sodium. As a matter of fact, the exceedingly low sodium in most cells gave rise to an early concept that the membrane must be relatively impermeable to this ion. However, studies using radioisotopes have shown conclusively that sodium readily penetrates the membrane. The low protoplasmic concentration appears not to be caused by inability of sodium to enter, but rather by some process that extrudes, or pumps

out, the sodium as quickly as it enters. Accordingly, the concept of a "sodium pump" has developed. Just how this sodium pump works is not known.

Donnan Equilibrium

One explanation of unequal ionic distribution lies in the so-called Donnan equilibrium, named after Donnan who published fundamental work in this field at the beginning of this century. It is sometimes referred to as the Gibbs-Donnan equilibrium, since it was Gibbs who first expressed the theoretical considerations that Donnan confirmed experimentally.

In brief, if a membrane separates two solutions of electrolytes in one of which there is a nondiffusible ion, an unequal distribution of the other, diffusible ions will result. This is because: 1) the sum of the equivalents of cations on either side of the membrane must be equal to the sum of equivalents of the anions on the same side, and 2) the products of the concentrations of any pair of diffusible ions on one side of the membrane must be equal to the product of the concentrations of the same pair of ions on the other side. Thus, if the nondiffusible ion has a negative charge, at equilibrium there will be a lower concentration of positive ion in the solution that does not contain the nondiffusible ion.

In Table 3.1 there are representative figures to illustrate the unequal ionic equilibrium that results from the presence of a nondiffusible ion. Consider the first set of figures in Example a. It can be seen that the positive ion, i.e., sodium, is certainly less concentrated, at equilibrium in Side 2 than it is in Side 1. Further it can be easily calculated that the sum of the cations equals the sum of the anions on each side. Thus, 133.3 equals 33.3 + 100; and 66.6 equals 66.6 + 0. Finally, the product of the concentrations of the diffusible ions on one side equals the product of those ions on the other side. Thus, 133.3 × 33.3 equals 66.6 × 66.6.

In Example b there are 2 cations present. At equilibrium they are both more concentrated in the side with the nondiffusible ion than they are on the other side. Again, at equilibrium the sum of the cations equals the anions on each side, and the product of any pair of diffusible ions equals the product of the same pair on the other side.

Without going into the mathematics involved, it should be obvious

TABLE 3.1. Initial and Final States of Systems Containing a Non-penetrating Ion X−, Having Two Equal Compartments Separated by a Membrane Permeable to the Other Ions. Concentrations in Arbitrary Units *

Example a

System	Initial State				At Equilibrium					
	Side 1		Side 2		Side 1			Side 2		
	Na+	X-	Na+	Cl-	Na+	Cl-	X-	Na+	Cl-	X-
One Cation 1	100	100	100	100	133.3	33.3	100	66.6	66.6	0
2	100	100	30	30	105.6	5.6	100	24.4	24.4	0
3	100	100	1	1	100.01	0.01	100	0.99	0.99	0
4	10	10	100	100	57.6	47.6	10	52.4	52.4	0

Example b

System	Initial State				At Equilibrium							
	Side 1		Side 2		Side 1				Side 2			
	Na+	X-	K+	Cl-	Na+	K+	Cl-	X-	Na+	K+	Cl-	X-
Two Cations 1	100	100	100	100	66.6	66.6	33.3	100	33.3	33.3	66.6	0
2	100	100	10	10	91.7	9.2	0.8	100	8.3	0.8	9.2	0

* Adapted from Harris, E. J., "Transport and Accumulation in Biological Systems," London, England, Butterworths Scientific Publications, 1956.

that the higher the concentration of the nondiffusible ion, the greater will be the resulting ionic imbalance. The protein in the cytoplasm acts as a nondiffusible ion and thus accounts, in part, for observed ionic inequalities. However, in most cells, the ionic imbalance is far greater than can be accounted for by calculation of the Donnan equilibrium; therefore, other factors must play important roles.

Electrochemical Interrelationships

The movement of ions through membranes is determined by: 1) the diameter of the hydrated ion, 2) concentration gradients, 3) osmotic pressures, and 4) electrical attraction by ions in the membrane and in the solutions on either side of the membrane. All of these factors are included in the term, electrochemical. If all the forces, that is, concentration, osmotic, and electrical, are known, then before movement takes place, the total force that will move that ion may be calculated. Such force is termed the electrochemical potential. The difference in the electrochemical potential between the two compartments determines the work required to move an ion from one compartment to the other.

MOVEMENT THROUGH THE CELL MEMBRANE BY ACTIVE PROCESSES

Thus far, the movement of water and of solutes through the cell membrane has been considered only in the light of such forces as diffusion, osmosis, concentration gradients, and electrical charge. All of these factors combined have been termed the electrochemical gradient. In other words, substances will move through a membrane down the electrochemical gradient. But there are many examples of movement of substances in the opposite direction, that is "uphill," up or against the electrochemical gradient. Clearly, work must be done to go against the tide. The energy for this work is provided by the cell. The transfer against a gradient is termed **active transport,** or **secretion.**

Active Transport of Water

There is now considerable question as to whether or not water is actively transported. It has been the custom to use the kidney as a

prime example of the active transport of water. As is well known, the kidney elaborates urine that is hypertonic in reference to blood plasma. This was long considered to be due to the active transport of water from the renal tubules back into the blood, thereby concentrating the urine. To be sure, such movement of water does occur, but this movement is now thought to be due to osmotic attraction rather than to an active transport mechanism. It is possible, however, that active transport of water may occur by the process of pinocytosis to be explained below.

Active Transport of Monosaccharides

The monosaccharides are relatively small molecules that penetrate most membranes and move in accord with concentration gradients. There are, however, many examples in which the monosaccharides do not follow such gradients. The kidney is an excellent example. Normally there is no glucose in the urine. Yet it has been clearly shown that glucose readily filters through the renal glomeruli. In the plasma there is about 1 mg of glucose per ml. If 130 ml of plasma filter through the human glomeruli, this means that about 130 mg of glucose per minute enter the lumen of the tubules. Yet none reaches the final urine (Fig. 3.6). Obviously, all of it must be reabsorbed by the tubule cells. Such complete removal runs against the gradient and can occur only by virtue of active transport. It is possible to treat the kidney with a drug that "poisons" the responsible metabolic process. When this is done, no glucose is reabsorbed by the tubule cells.

There are other examples of the active transport of monosaccharides. In the small intestine it has been demonstrated that the various monosaccharides are absorbed by the intestinal cells independent of their concentration gradient. Again, the mechanism is not known.

Active Transport of Ions

The most carefully documented example of active transport involves the ions, especially potassium and sodium. As has been mentioned, there is a striking imbalance in the ionic distribution between the inside and the outside of the cell. This imbalance cannot be explained by the Donnan equilibrium effect. Beyond question some active cellular process must exist to maintain the imbalance. The

use of radioisotopes has proved that this is not a static mechanism. In other words, the ions are not simply fixed, being incapable of moving through the membrane, but rather the ions move through the membrane freely. There seems to be a constant inward and outward flux of the various ions. Yet in the living cell the imbalance is maintained. Under certain conditions, such as the propagation of an

PLASMA GLUCOSE
I mg/ml

GLOMERULUS

TUBULE CELLS

FILTRATE
130 ml/min

FLUID REABSORPTION
129 ml/min

GLUCOSE
130 mg/min

GLUCOSE REABSORPTION
130 mg/min

URINE   I ml/min
NO GLUCOSE
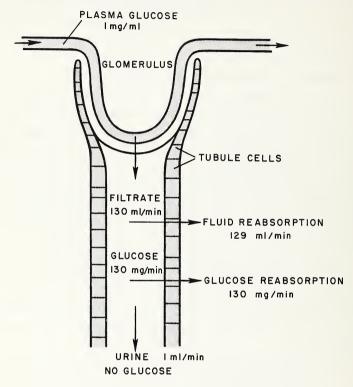
Fig. 3.6.   Active Transport of Glucose by the Kidney Tubule Cells.

impulse by a nerve cell, there is a rapid inward movement of sodium and an outward movement of potassium. But then, the reverse occurs, the sodium that has entered is pumped out and potassium reenters to once again reestablish the ionic imbalance. This extrusion of sodium is obviously an active process and to describe it, but hardly to explain it, the expression, sodium pump, has been coined. Whether or not there is a "pump" for potassium and other ions is not known, and the term does very little to clarify the mechanism.

Mechanism of Active Transport

The mechanism of active transport is almost wholly unknown. However, it has been amply demonstrated that interference with cell metabolism reduces or abolishes it. Thus, if the cell is denied oxygen, active transport soon ceases. Likewise there are many so-called metabolic poisons which block certain metabolic processes and, in this way, interfere with specific examples of active transport. Finally, it has been shown that certain enzymes or hormones control the active transport of some substances.

Elaborate working hypotheses have been suggested which purport to explain active transport. In general they suggest complex chemical reactions which incorporate the substance to be transported at one cell border into a compound. Because of this incorporation, by simple diffusion the substance moves into the cell (Fig. 3.7). The compound then makes its way across the cell and at the other cell border decomposition takes place and the original substance is set free to diffuse across the membrane.
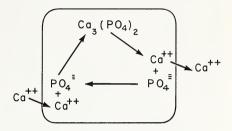
Fig. 3.7. A Possible Mechanism of Active Transport. The calcium ions diffuse into the cell down a concentration gradient. The calcium ions are taken out of solution by the formation of $Ca_3(PO_4)_2$. When this compound dissociates at the opposite membrane, calcium ions diffuse out of the cell. As a result, the concentration of calcium ions on the right of the cell is greater than on the left.

This idea of the incorporation of a substance into a compound and ultimately setting it free is attractive because at each border there is a concentration gradient down which the substance can move to be transported through the membrane. Yet, the total effect is to move the substance through the entire cell against a gradient.

The above general explanation is plausible for active transport through a cell, but it must be stretched somewhat thin to account for the active transport of a substance into or out of a cell across only one membrane. There could be similar compound formation within the cell membrane itself, but very little evidence for this possibility currently exists. Of course, this entire problem is somewhat confused

because in many types of cells it is difficult to say just where the cell membrane ends and the cytoplasm begins. There is the growing belief that the so-called cortex of the cell may play an important role in active transport. Again it is postulated that the substance undergoing movement is taken up or bound in some manner by the cortex. This would cause the substance to move through the membrane. There then must be some mechanism by which the substance is released by the cortex to the interior of the cell.

There is good evidence that some cells at least are capable of moving water by a process termed **pinocytosis,** which means "a cell that drinks." In this process there is a discontinuous uptake of fluid, in the form of droplets, by the cell. The droplet seems to come in contact with the cell membrane which then forms a valley into which the droplet moves. The upper edges of the membrane then come together, fuse, and thus seal off a vesicle. Now the droplet is completely surrounded by a membrane which lies within the cell (Fig. 3.8). This membrane then disintegrates leaving the droplet within the cytoplasm.
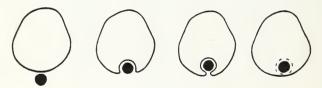
Fig. 3.8.   Pinocytosis. The droplet is engulfed by the cell and then ultimately comes to rest within the cell.

A similar process involving solutes has been termed **cytopemphis,** which means a "cell pustule." In cytopemphis the particle is transported completely through the cell. In other words, it is first engulfed and moved into the cell as outlined for pinocytosis. But then the reverse procedure occurs at the opposite membrane and the particle is discharged from the cell. It is possible that water may be moved through the cell in the same manner. This mechanism differs from phagocytosis in that the particle is transported through the cell rather than held or digested by the cell.

Whether cells in general are capable of transporting water and

solutes by pinocytosis and cytopemphis is not yet known. But if this proves to be the case then the significance is great for herein lies an explanation for active transport, as well as for the transport of substances composed of large molecules.

FACTORS THAT ALTER CELL MEMBRANE PERMEABILITY

The permeability of the living cell membrane varies under different conditions. Marked permeability changes are evoked by severe injury or death of the cell. A very common example of the influence of injury on permeability is the swelling and discoloration of a part of the body that has been subjected to trauma. The trauma alters capillary permeability; plasma and some erythrocytes pour out into the tissue spaces, thus causing the swelling. Alteration of the hemoglobin in the erythrocytes causes the discoloration. But factors other than injury alter permeability. These will be briefly considered.

Ionic Effect on Permeability

The ionic environment of the cell membrane strikingly alters its permeability. It has long been known that the addition of sodium or potassium ions to the surrounding medium increases permeability. In contradistinction, the addition of calcium or magnesium ions decreases it. The situation, however, is not that clean-cut, or simple. For example, if the cell is suspended in an isotonic, non-electrolyte solution, sodium and calcium have the same effect, namely to decrease the permeability of the cell to water. On the other hand, these have opposite effects when the cell is suspended in an isotonic salt solution. The complexity of the problem is further shown by the fact that the various ions exert an influence which apparently depends upon the percentages of other ions present.

The mechanism of the ionic effect is in the realm of hypothesis. It has been suggested that the membrane may undergo phase change. That is, under one set of conditions it may be an oil-in-water emulsion. With the water being continuous, permeability of water-soluble substances would be increased. A change to a water-in-oil emulsion would decrease permeability. Another suggestion is that the various ions alter the chemical composition of the membrane and thus change

the partition coefficients for various substances. Neither suggestion explains all of the observed ionic effects, nor is there convincing evidence to support either one.

Influence of Narcotics

A narcotic is a substance that decreases sensibility. Cells treated with narcotics undergo reversible loss of irritability, that is, they do not respond to the usual stimuli. It has long been thought that the explanation must lie in the alteration of the cell membrane by the narcotic agent so as to render it completely impermeable. But more recent studies, especially those employing radioisotopes, have shown that although the permeability of the membrane may, indeed, be altered, that alteration is not always one of decreased permeability. In some cases permeability is increased. Thus, it can be said that narcotics do alter permeability, but in either direction. Just how such changes in membrane permeability result in decreased cell responsiveness awaits further insight into the mechanism of irritability.

Influence of Temperature

There is general agreement that the higher the temperature the greater the transport rate. Of course, if the temperature is raised beyond a critical point, irreversible changes occur. The increased rate of movement of substances through the cell membrane as the temperature is raised cannot be accounted for solely on the basis of increased molecular activity. There is apparently an alteration in the cell membrane.

There are many environmental factors that alter cell membrane permeability. These include radiation, electric currents, pH, and a long list of inorganic and organic substances. But until greater knowledge of the mechanism of membrane permeability is obtained, there is little point in simply listing the effects that these various environmental factors have on permeability.

SUMMARY

The limiting structure of the plant cell is termed the **cell wall;** in animal cells, the **plasma membrane.** Additional layers are sometimes present for protection. The functional cell membrane is very

thin. In the erythrocyte most measurements disclose a cell thickness of about 150 Å, but in the living unaltered state it may be thicker.

The cell membrane is composed of **protein, lipid,** and various salts, primarily **calcium.** The fat is thought to be arranged in a double layer and to be covered on each surface by a very thin layer of protein. **Pores** are assumed to traverse the membrane.

The surface of most cells has a negative charge. The surface tension is quite low, being of the order of 1 dyne per centimeter, or less.

The forces that move water through the membrane are: 1) **molecular activity,** 2) **hydrostatic pressure,** 3) **osmotic pressure,** and 4) **active transport.** The unit used to express permeability to water is the number of cubic microns of water entering the cell per minute per unit area of membrane per atmosphere of difference of osmotic pressure between the internal and external media. Using these units, permeability to water has been found to vary from 0.2 to 2.8 cubic microns per minute.

The factors that determine the movement of solutes through the cell membrane are: 1) **molecular size,** 2) **partition coefficient,** 3) **concentration gradient,** 4) **charge,** and 5) **active transport.**

A substance may move through a membrane by **active transport** in a direction opposite to the **electrochemical gradient.** The mechanism of active transport is not understood. It is generally postulated that there is some specific chemical reaction which supplies the necessary energy. **Pinocytosis** and **cytopemphis** may play roles.

Cell membrane permeability is altered by changes in the environment of the cell. These include: 1) **ionic concentration,** 2) presence of **narcotics,** 3) **temperature,** 4) **radiation,** 5) **electric currents,** and 6) **pH.**

Problems

1. Taking into consideration current concepts of cell membrane structure and composition, explain the role of pore size and partition coefficients in the movement of solutes through the membrane.
2. Differentiate between a semipermeable and a selectively permeable membrane.
3. What role do cytoplasmic proteins play in the movement of water through the cell membrane?
4. Discuss the movement of ions through the cell membrane in the light of atomic size, charge, and the Donnan equilibrium.

5. Define:

  a. Isotonic

  b. Electrochemical gradient

  c. Permeability unit

  d. Partition coefficient

  e. Pinocytosis

  f. Surface tension

Additional Reading

Clark, H. T., ed., "Ion Transport Across Membranes," Academic Press, New York, N. Y., 1954.

Davson, H. and Danielli, J. F., "The Permeability of Natural Membranes," The Macmillan Co., New York, N. Y., 1943.

De Robertis, E. D. P., Nowinski, W. W., and Saez, F. A., "General Cytology, 3rd ed., W. B. Saunders Co., Philadelphia, Pa., 1960.

Harris, E. J., "Transport and Accumulation in Biological Systems," Butterworths Scientific Publications, London, England, 1956.

Murphy, Q. R., ed., "Metabolic Aspects of Transport Across Cell Membranes," University of Wisconsin Press, Madison, Wisconsin, 1957.

PART II

Intracellular Activities

CHAPTER 4

METABOLISM: BASIC PRINCIPLES

THE AGGREGATE of all the physical-chemical processes constantly taking place in the living cell is termed metabolism. These processes include those which utilize energy to synthesize complex substances from simpler elements as well as the reactions which release energy. Processes which result in biosynthesis are embraced by the term **anabolism. Catabolism** includes all reactions in which complex molecules are converted to smaller ones, generally with the release of energy.

ENZYMES

The rate of chemical reaction may be altered by the presence of substances known as catalysts. **Catalysts** do not enter into the reaction in the sense that they are used up, or converted, into another form. Instead a catalyst alters the rate of a chemical reaction and may be recovered intact when the reaction is finished. Most commonly the term is restricted to an agent that accelerates a chemical reaction, although strictly speaking there are substances, also termed catalysts, which inhibit chemical processes.

An **enzyme** is a catalyst formed by living cells and thus is sometimes termed a biocatalyst. Insofar as is known, all enzymes are proteins and thus have a high molecular weight.

Types of Enzymes

The main chemical compound which undergoes alteration in a chemical reaction upon which the enzyme exerts its influence is

termed the **substrate.** Enzymes are often classified and named according to the substrate upon which they act. For example, if an enzyme acts on a phosphate compound it may be termed a phosphatase. The suffix -ase is simply added to the specific name of the substrate. Another way to classify enzymes is according to the class of subtances upon which they act. In this classification, the suffix -lytic is used to name the type of enzyme. For example, enzymes that act on lipids are termed lipolytic. These simple systems for naming enzymes are followed, for the most part, but there are exceptions. Some enzymes that have been long recognized were named prior to the general adoption of these suffixes and retain names that were given to them somewhat arbitrarily. One such example is ptyalin, an enzyme found in saliva. Under the more modern system of nomenclature this same enzyme is also known as amylase since amyl- is a combining form denoting "pertaining to starch."

In addition to classifying enzymes according to the substrate or class of substances upon which they act, there is a rather recent trend to classify them according to their chemical composition. But since so little is known concerning the chemical composition of the vast majority of enzymes this system is still very limited.

Protein Nature of Enzymes

As previously mentioned, it is commonly believed that all enzymes are proteins. Until very recently the thought persisted that the active agent itself was merely combined with, or in someway bound to the protein molecule. However, enzymes have now been prepared in the crystalline state. These preparations retain full potency and only lose their potency when the protein is denatured. The evidence is therefore quite convincing that it is the protein itself which is responsible for the catalytic action. Enzymes exhibit all the classical properties of proteins, namely high molecular weight, colloidal behavior, slow diffusion, inability to pass through most living membranes, and movement in response to an electrical current.

Factors that Alter Enzyme Action

Temperature, pH, substrate and enzyme concentration, and the presence of other chemical substances may all modify enzymatic activity. Usually, an elevation in temperature accelerates the action of

an enzyme. It has been found that a rise of 10°C hastens the reaction two or three times. However, there is a limit to the acceleration of enzymes by heat because above a certain temperature the enzyme is inactivated. Thus, if one plots the rate of reaction against increasing temperature, it is found that at first the rate increases and then decreases (Fig. 4.1). The increase is due to two factors: 1) the influence of heat on the chemical reaction itself and 2) increased enzyme activity. The decrease results from inactivation of the enzyme by heat. There is thus an optimum temperature for each enzyme.
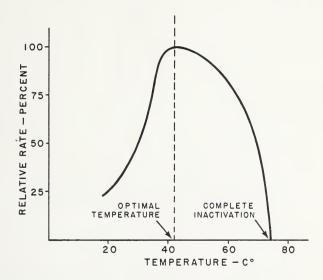
Fig. 4.1.   The Influence of Temperature on Enzyme Action.

A similar graph is obtained when rate of reaction is plotted against the acidity, that is, the pH of the medium (Fig. 4.2). There is thus an optimum pH as well as an optimum temperature for each enzyme.

In most chemical reactions an increase in concentration of the reactants increases the rate of the reaction. The same is true for processes enhanced by enzymatic action. Thus the rate of reaction may be increased by raising the concentration of the substrate, of the enzyme, or of both.

Until quite recently the mammalian physiologist avoided studying the cell, as though it were an area outside his jurisdiction. The same was true of the pharmacologist. Thus, hormones and drugs and other

substances were analyzed in terms of various systems, such as circulation or respiration or excretion, but little or nothing was said about the intracellular response. Such avoidance of the cell is no longer possible, and for the final understanding of the action of these various substances it is necessary to enter the cell, so to speak, to comprehend the various intracellular processes, the influence of enzymes on these processes, and then to analyze the alterations evoked by the substance
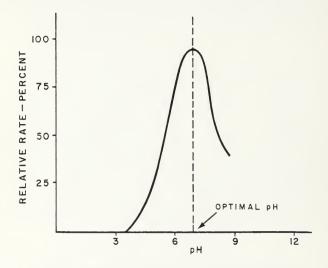
Fig. 4.2.   The Influence of Acidity on Enzyme Action.

in question. It has been found, when this is done, that the activity of many hormones and drugs results from either an enhancement or inhibition of enzyme action. In some instances it is thought that the concentration of the enzyme is varied and that in this way the intracellular process is altered. In view of the fact that perhaps every intracellular process is enzyme-controlled it is obvious that this subject is of broad and vital importance.

Enzyme Specificity

It is generally true that enzymatic activity is restricted to a specific substrate. Thus, lipase will hasten the digestion of lipids, but exerts no influence on the catabolism of protein or carbohydrate. This property of enzyme action is referred to as enzyme specificity.

Mechanism of Enzyme Action

Enzymes may influence the rate of biochemical reactions in several ways. There is evidence to indicate that in some instances the enzyme reacts with the substrate to form a new compound. This intermediate compound then undergoes a further reaction which results in the liberation of the enzyme in its original form but the substrate in an altered form. Thus, the over-all change is in the substrate alone.

Another way in which enzymes function is by lowering the energy of activation of a reaction. The **energy of activation** is a term which has been applied to the minimal amount of energy required of a molecule to take part in a reaction. Not all molecules of the same substance have the same amount of kinetic energy. Thus, the rate of a chemical reaction may well depend upon the concentration of high energy molecules. It has long been known that temperature increases the concentration of high energy molecules, and this is thought to be the explanation for the increase in the rate of chemical reactions at higher temperatures. Likewise, catalysts may exert their enhancing influence by somehow lowering the threshold, that is, the minimal amount of energy required for molecules to take part in the reaction. At this lower threshold more molecules can take part than at the higher threshold, and therefore the reaction proceeds more rapidly.

This explanation of the mechanism of enzymatic action raises more questions than it answers. One wonders why some molecules have more energy than others and, more importantly, how an enzyme makes it possible for a reaction to proceed at a lower energy level?

Whatever the answers to these questions may be, the fact remains that the energy of activation has been measured for various reactions and in every case, in the presence of a catalyst that speeds the reaction, this energy has been found to be lowered. For example, the hydrolysis of sucrose, with only the hydrogen ion as the catalyst, has an energy of activation of about 25,000 calories per mole; when yeast invertase is added, the energy of activation is less than 10,000. Even more striking is the case of decomposition of hydrogen peroxide. Without a catalyst this reaction has an energy of activation of some 18,000 calories per mole. When the enzyme catalase is present, it is less than 2,000.

ENERGY METABOLISM

Energy may be defined as the capacity for doing work. As has already been emphasized, the properties characteristic of the living cell involve work and therefore require energy. For example, transport across an inanimate membrane may occur by simple diffusion or by filtration. The energy for such transport is not provided by the membrane. In the living cell, however, transport across the cell membrane may occur against an electrochemical gradient. Clearly, work is done and here the energy for the work is supplied by the cell. The same is true of such cellular activities as respiration and irritability, growth and reproduction. The energy for these activities must be derived from various intracellular chemical transformations. As indicated, these chemical processes are transformations, that is to say, the energy is not created but simply converted into a form usable for specific work. The source of this energy is food in the form of nutrients that can enter the cell. Since energy is constantly utilized and much of it is lost in the form of heat, for survival there must be a continuing source of nutrients.

Cellular Oxygen Consumption

It is possible to measure the oxygen consumption of cells. This is most often done by manometric methods. The rate, known as the **respiratory rate,** is generally expressed as the volume of oxygen consumed per unit weight of the tissue under study. The oxygen volume is always corrected to standard temperature and pressure. The symbol used to denote the respiratory rate is Q_{O_2}.

Unicellular organisms may be studied in this way, or tissues from multicellular forms may be used. Not only does the respiratory rate vary from one unicellular organism to another, but there is a wide difference in the rate of oxygen consumption between tissues taken from the same animal. A few examples to illustrate this point are given in Table 4.1.

The oxygen consumption varies with the temperature. In general, a rise in temperature increases the respiratory rate, and a fall decreases it. This is thought to be due to the effect of temperature on the intracellular oxidative processes. On the other hand, the oxygen concentration in the environment of the cell does not influence the

rate of consumption until the concentration falls to very low levels. In the air, and in the environment surrounding the cells in multicellular organisms, the oxygen concentration is usually much higher than this minimal level. Finally, there are many substances that either enhance or inhibit the respiratory rate. As a matter of fact, the influence of various drugs on the living organism depends upon the effect these compounds have on the intracellular processes.

TABLE 4.1. Representative Respiratory Rates. Q_{O_2} Expressed in Ml Per Gram Per Hour

Type of Cell	Q_{O_2}
Paramecium caudatum	1.0 (wet weight)
Amoeba proteum	0.2 (wet weight)
Bacillus fluorescens	4100.0 (wet weight)
Rat kidney	31.0 (dry weight)
Rat thyroid	13.0 (dry weight)
Rat heart	5.0 (dry weight)
Rat skin	0.8 (dry weight)

The Calorie

Energy in the multicellular form is often expressed in calories. There are two units of heat, the small calorie and the large Calorie, written with a capital C. The small calorie is defined as the amount of heat necessary to raise the temperature of 1 gm of water from 14.5 to 15.5°C. The large Calorie is one thousand times greater. By definition, it is the amount of heat necessary to raise the temperature of 1 kg of water from 14.5 to 15.5°C. This is the unit most commonly used to express food values.

Food Values

When the various foodstuffs are burned in an atmosphere of oxygen, heat is produced. If 1 gm of each of the three foods is placed in a chamber, a so-called **bomb calorimeter** (Fig. 4.3), and ignited, the following values are obtained:

Protein	5.3 Calories
Carbohydrate	4.3 Calories
Fat	9.5 Calories

When 1 gm of each of these same substances is burned in the living organism, the following values are noted:

Protein   4.1 Calories
Carbohydrate   4.1 Calories
Fat   9.3 Calories

The values for carbohydrate and fat are remarkably similar whether oxidized in the calorimeter or in the living organism. Protein differs, however, and this difference is thought to be due to incomplete utilization by the organism. It must be emphasized that the above values are obtained in multicellular organisms. It is true that the ultimate utilization of each foodstuff occurs intracellularly, but before this can take place the complex molecules are altered by the digestive processes until the basic compounds, that is, amino acids, monosaccharides, glycerol and fatty acids, are made available to the cell.
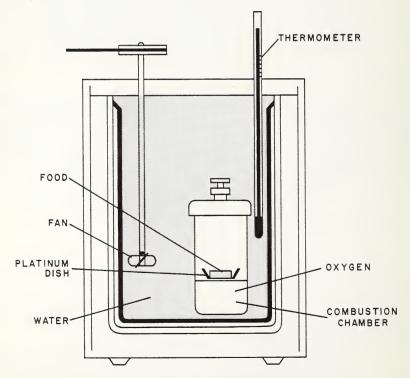
Fig. 4.3. The Bomb Calorimeter. The food is ignited in an atmosphere of oxygen. The resulting heat elevates the temperature of the water.

Respiratory Quotient

One of the classic early demonstrations was that a candle would not burn, and life could not be sustained, in an atmosphere devoid of oxygen. It has been shown further that whether foods are burned in a bomb calorimeter or in the living organism a specific quantity of oxygen is utilized for each type of food and, as a consequence, a definite amount of carbon dioxide is produced. In other words, it can be concluded that energy and metabolic heat result from the oxidation of various nutrients. Oxygen is utilized and carbon dioxide produced, but this is not always a molecule for molecule relationship. The ratio of carbon dioxide to oxygen is termed the respiratory quotient ($R.Q.$), that is:

$$R.Q. = \frac{CO_2}{O_2}$$

If glucose is oxidized the reaction is:

$$C_6H_{12}O_6 + 6O_2 = 6CO_2 + 6H_2O$$

It is seen that there are 6 molecules of oxygen used and the same number of carbon dioxide molecules produced. The $R.Q.$ when glucose is oxidized is unity, i.e., 1.

It is found that the oxidation of body lipids results in an $R.Q.$ of only 0.71. The $R.Q.$ for the oxidation of protein, an extremely complex process, has been estimated to be about 0.8. Finally, on a mixed diet, the organism as a whole exhibits an $R.Q.$ of about 0.82. This means that the three foodstuffs are being oxidized simultaneously. If the $R.Q.$ is found to be close to 1, it is interpreted to mean that a high percentage of carbohydrate is being utilized. It is possible to exceed 1.0 in experimental animals which are given only large amounts of carbohydrate to eat. In such cases there is conversion of the carbohydrate to fat. Since carbohydrate has more oxygen than fat in its molecule, the conversion makes oxygen available and thus the oxygen consumption of the animal during this conversion falls below the carbon dioxide consumption and the $R.Q.$ exceeds 1.0.

Calorimetry

It has been seen that the ultimate source of energy is the oxidation of nutrients. In this process a definite quantity of oxygen is utilized

and a specific amount of heat produced. Thus, an indication of the rate of metabolic processes may be obtained by either measuring the heat produced or the oxygen utilized. Total heat production is determined in what is termed **direct** calorimetry. **Indirect** calorimetry involves the measurement of oxygen utilization and from this value heat production is calculated.

Because the direct method necessitates bulky, expensive equipment in which the entire organism must be placed in order to determine the total heat production, the indirect method is more often used. In this latter procedure it is only necessary to determine the quantity of oxygen utilized by the organism per unit time. For both man and lower animals simple equipment is available to permit this measurement.

Again consider the reaction:

$$C_6H_{12}O_6 + 6O_2 = 6CO_2 + 6H_2O$$

One gm-mole weight of glucose is 180 gm. According to the above equation this quantity must react with 192 gm of O_2. Avogadro's principle states that 1 gm-mole of any gas has a volume of 22.4 liters. Therefore 6 gm-moles will occupy 134.4 liters (6 × 22.4). In other words, 134.4 liters of oxygen react with 180 gm of glucose. Therefore, 0.75 liter of O_2 must react with each gram of glucose (134.4/180). It has been shown by direct calorimetry that each gram of glucose oxidized in the organism produces 4.1 Calories. And it has been calculated above that 1 gram of glucose while being burned utilizes 0.75 liter of O_2. One liter of oxygen, then, in reaction with carbohydrate must produce 5.47 Calories (4.1/0.75). Similar values have been calculated for fat and protein (Table 4.2).

TABLE 4.2. Energy Relationships

	Carbohydrate	Fat	Protein
Liters of O_2 per gram	0.75	2.03	0.97
Liters of CO_2 per gram	0.75	1.44	0.78
R.Q.	1.00	0.71	0.80
Calories per gram	4.10	9.30	4.10
Caloric value of 1 liter O_2	5.47	4.58	4.23

From Table 4.2 it is obvious that the caloric equivalent of a liter of oxygen depends upon the type of food being utilized. In indirect calorimetry it is assumed that all three are burned. The average $R.Q.$ found under these circumstances is 0.82. At this $R.Q.$ the caloric equivalent is 4.825. Accordingly, once the quantity of oxygen utilized has been determined and corrected for standard conditions, one need only multiply it by 4.825 to ascertain the Calories produced by the organism in unit time. It is the custom to express this value in Calories per square meter of body surface per hour. In the adult human male this value is about 40.

INTRACELLULAR OXIDATION

The ultimate source of biological energy is intracellular oxidation. Clearly, the mechanism by which oxidation occurs and the resultant energy utilized is of primary importance.

The equation representing the oxidation of glucose has been presented above. This appears to be a simple reaction that might be thought to proceed easily. It should be recalled, however, that glucose is kept in the home as table sugar, and in the laboratory for years in the presence of an abundance of oxygen and yet very little oxidation occurs. One way of causing this reaction to proceed is by greatly elevating the temperature. In the cell, quite obviously, another method must be utilized.

Heat can be utilized by various types of machines to do work. But since the temperature of the mammalian cell does not vary appreciably, heat resulting from oxidation must be dissipated, rather than utilized for biological work. In the cell there are thought to be specific reactions for each process requiring work. For example, in discussing the mechanism of active transport (cf. page 55), a process in which work is required, it was suggested that the substance being transported is synthesized, by the cell, into a new compound. This new compound then makes its way across the cell where it undergoes breakdown to liberate the original substance which then diffuses out of the cell. The formation and breakdown of such a compound does work in transporting the substance against a gradient. By the same token, the formation and breakdown requires energy. This energy is made available by some other reaction, such as the oxidation of

nutrients. Thus, one visualizes the cell as containing numerous interlocking reactions, the energy from one being utilized in order for another to proceed. In this way energy can be transferred repeatedly and yet the temperature of the cell is maintained constant.

It is the role of the cell physiologist and biochemist to elucidate these intracellular reactions. In this section, the ultimate source of energy for these reactions, namely oxidation, will be discussed.

Oxidation-Reduction Reactions

Any reaction in which oxygen takes part is termed oxidation. However, this definition is too narrow. In the broader definition, oxidation occurs whenever there is the loss of an electron. In contradistinction, the acquisition of an electron is termed reduction. Consider the reaction:

$$2H_2O \longrightarrow 4H+ + 2O^{--}$$

It is seen that each atom of hydrogen has lost an electron and therefore has a positive charge, while each atom of oxygen has gained two electrons, and thus has a negative charge. According to the above definition, hydrogen has been oxidized and the oxygen reduced.

By similar reasoning, the following reactions represent oxidation:

$$Co++ \longrightarrow Co+++ + 1 \text{ electron}$$
$$Fe++ \longrightarrow Fe+++ + 1 \text{ electron}$$

The reverse reactions would occur by reduction.

In the case of organic compounds, oxidation and reduction is somewhat more complex. There is in such reactions the loss of a hydrogen ion, i.e., a proton. In most such reactions there is also the loss of an electron. At times these losses occur simultaneously, at others sequentially. For example:

$$
\begin{array}{ccc}
\overset{\displaystyle H}{\underset{\displaystyle OH}{CH_3 - C - H}} \longrightarrow &
\overset{\displaystyle H}{\underset{\displaystyle OH+}{CH_3 - C - H+}} + 2 \text{ electrons} \longrightarrow &
\overset{\displaystyle H}{CH_3 - C = O} + 2H
\end{array}
$$

In this reaction ethanol has been oxidized. Two electrons and two hydrogen atoms have been lost. Thus the definition of oxidation is made still broader to include **dehydrogenation.** In the oxidation of a compound in this way, the compound that gives up hydrogen is

termed the **hydrogen donor.** The compound that incorporates the hydrogen is called the **hydrogen acceptor.**

Because under appropriate conditions oxygen avidly accepts electrons and hydrogen ions, it is an excellent oxidant. Within the cell there are many other oxidants, but ultimately oxygen is utilized. It is for this reason that the determination of oxygen consumption is a measure of the total metabolic activity of the cell.

Redox Potentials

The term "redox" is a generally used abbreviation for oxidation-reduction. Oxidation and reduction, as explained above, involve the transfer of electrons. The relative ease with which this transfer is made is indicated by the potential of the system, that is, the redox potential expressed in volts.

If a metal bar is immersed in an aqueous solution or in pure water some of the atoms of that metal will give up electrons and thus become ions. For example, if the metal is iron, there will be iron ions, that is Fe^{++} in solution at the surface of the metal. The electrons remain on the metal, therefore it has a negative charge. The iron ions in solution have a positive charge. Since opposite charges attract one another, there will be recombining of electrons and ions to reform iron atoms. At equilibrium the formation of ions and the formation of atoms proceed at the same rate. Since there will always be some ions in solution, the metal will always be negative and therefore there is a potential difference between the metal and the layer of ions. This potential difference is characteristic of each metal in a particular solution and is termed the **electrode potential.**

The electrode potential varies with the substance used for the electrode. Therefore if two such systems are connected by a wire, a current will flow. This current may be measured by a galvanometer. However, it is the practice to use a potentiometer in which a known electromotive force (emf) is adjusted to prevent current flow, that is, it is adjusted until the galvanometer reads zero. This emf, expressed in volts, measures the difference between the two electrode potentials. In order to have a reference point it is the agreed custom to express redox potentials relative to a **normal hydrogen electrode.** Such an electrode is made by bubbling hydrogen over a piece of platinum which has been covered with very fine platinum particles,

so-called platinum black. Hydrogen is adsorbed by the platinum and an equilibrium between hydrogen atoms and ions established. The term "normal" refers to the solution in which this electrode is immersed, namely a $1N$ H^+ solution.

Using the hydrogen electrode as a reference, all other types of electrodes may be tested against it and the voltage necessary to prevent current flow measured. Since the hydrogen electrode is arbitrarily assigned a zero value, the redox potentials of other systems are found to vary on both sides of zero, that is, plus and minus. For example, in reference to the normal hydrogen electrode at a temperature of 25°C, a copper electrode is rated +0.34, a sodium electrode −2.71. This means that when the hydrogen and copper electrodes are connected by a wire, electrons move from the hydrogen electrode to the copper electrode. To prevent such movement emf must be added, thus the positive sign. Insofar as oxidation-reduction systems are concerned, it means that copper accepts electrons more avidly than does hydrogen and thus is a stronger oxidizing agent. Conversely, hydrogen is the stronger reducing agent. Sodium, which has a negative potential relative to hydrogen, must be a still stronger reducing agent than either copper or hydrogen. From this discussion it can be seen that the redox potential gives an indication of the relative strength of the substance in terms of oxidation and reduction—the more positive the potential the stronger the ozidizing power, the more negative the stronger the reducing power.

If substances are arranged according to their redox potential ranging from the most negative to the most positive, a series results which is called the **electromotive series** (Table 4.3). All of these values are not only relative to hydrogen, but to each other as well. Consider zinc and iron. Under the conditions outlined above zinc shows a potential of −0.74 and iron −0.44. In relation to hydrogen both are stronger reducing agents. But it also means that iron in relation to zinc is an oxidizing agent. To put it another way, iron is capable of oxidizing zinc, but it reduces hydrogen.

Redox Potentials of Biological Systems

The redox potential of compounds normally found within cells may also be determined. The method is essentially the same, and comparison is made against the hydrogen electrode. However, since

most biological substances cannot serve as electrodes, a platinum electrode, which is chemically inert, is placed in a solution of the substance being determined. The potential of such a solution depends upon the percent oxidation or reduction. That is to say, when it is completely reduced, it will have one value; when completely oxidized, another. Thus, known amounts of an oxidizing agent can be added to a substance and the potential measured for each mixture. The plotted results of this procedure using a biologically important substance, cytochrome, are shown in Fig. 4.4. In practice the procedure is simplified by using the following equation:

$$E = E_o + \frac{0.06}{n} \log \frac{\text{(oxidant)}}{\text{(reductant)}}$$

where: E = redox potential in volts
E_o = redox potential of the half-reduced system, i.e., oxidant/reductant = 1
n = number of electrons transferred per mole

Thus one need only determine E_o after which E for any degree of reduction may be calculated.

TABLE 4.3. Electromotive Series. The Redox Potentials, Expressed in Volts, Are Compared with a Normal Hydrogen Electrode at 25°C. The Sign of the Potential is Arbitrary.

Electrode Reaction	Redox Potential
$Au = Au^{+++} + 3$	+1.5
$2Cl^- = Cl_2 + 2e$	+1.4
$Hg = Hg^{++} + 2e$	+0.86
$Ag = Ag^+ + e$	+0.8
$Fe^{++} = Fe^{+++} + e$	+0.74
$Cu = Cu^+ + e$	+0.51
$Cu = Cu^{++} + 2e$	+0.34
$H_2 = 2H^+ + 2e$	0.00
$Pb = Pb^{++} + 2e$	−0.12
$Ni = Ni^{++} + 2e$	−0.22
$Fe = Fe^{++} + 2e$	−0.44
$Zn = Zn^{++} + 2e$	−0.74
$Mn = Mn^{++} + 2e$	−1.1
$Al = Al^{+++} + 3e$	−1.7
$Mg = Mg^{++} + 2e$	−2.4
$Na = Na^+ + e$	−2.7
$K = K^+ + e$	−2.9

It should be noted that the above equation only holds when the temperature is 30°C. Furthermore, since the pH of the system plays a role in the redox potential, both the pH and the temperature must be known in order to compare systems. The reference hydrogen electrode is always 1 normal, but this is a high acidity which never exists in biological systems; thus for meaningful values the solution being examined is usually kept at a pH similar to that normally
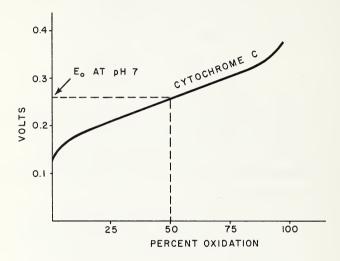
Fig. 4.4. Potentiometric Titration Curve for Cytochrome C. The potential difference is between the standard hydrogen electrode and the cyctochrome solution at pH 7 and at a temperature of 30°C. E_0 is equal to the potential difference when the cytochrome solution contains equimolar amounts of the oxidized and reduced forms.

found intracellularly, that is, pH = 7.0. To illustrate the alterations evoked in the redox potential by pH it need only be pointed out that a hydrogen electrode in a solution of pH 7 compared with the normal hydrogen electrode shows a potential of about −0.42 volts.

Use of Redox Potentials

It was pointed out above that metals may be arranged in an electromotive series and from this series the relative reducing or oxidizing power of each metal determined. The same may be done with biological redox systems. Table 4.4 lists some of these systems. Any

system is capable of oxidizing all systems of lower potential, and reducing a system of higher potential. Thus, a knowledge of the redox potentials of various biological systems permits an explanation as to why reactions proceed in a particular way.

TABLE 4.4. Biological Oxidation-Reduction Systems. All Values, Expressed in Volts, Are at pH 7

System	Redox Potential
$H_2O/\frac{1}{2}O_2$	0.82
$H_2O_2/\frac{1}{2}O_2$	0.30
Cytochrome a Fe++/Fe+++	0.29
Cytochrome c Fe++/Fe+++	0.26
Butyryl CoA/crotonyl CoA	0.19
Hemoglobin/methemoglobin	0.17
Succinic acid/fumaric acid	0.03
Alanine/pyruvic acid + NH_4^+	−0.13
Glutamic acid/α-ketoglutaric acid + NH_4^+	−0.14
Lactic acid/pyruvic acid	−0.19
DPN+/DPNH + H+	−0.32
Malic acid/pyruvic acid + CO_2	−0.33
Xanthine/uric acid	−0.36
$\frac{1}{2}H_2/H+$	−0.42

Another use of redox potentials is to determine the free energy of a reaction. The following equation is used:

$$\Delta F^o = \frac{-nF\,\Delta E_o}{4.18}$$

where: ΔF^o = the free energy change of the reaction expressed in calories per mole

n = the number of electrons or hydrogens

F = the faraday which is equal to 96,500 coulombs

ΔE_o = the difference between the redox values of the reactants

4.18 = a constant used to convert the values to calories

The value obtained when the above equation is solved is usually negative, which means that in the reaction, energy, expressed in terms of heat, has been given up and thus made available for other cellular work.

It is always of interest to know whether or not the redox potential of any particular reaction is the same *in vivo* as it is *in vitro*. Thus, attempts have been made to measure redox potentials in the living cell. The most direct way is to place microelectrodes in the cell, but this procedure obviously drastically alters the cell and therefore it is questionable whether or not the obtained potential represents a normal value. Results have also been obtained using indicator dyes. The color change of each particular dye indicates a certain potential. Remarkably consistent results have been obtained using this method but still the objection persists that the dye itself alters the metabolic activity of the cell. In brief, there is no method that is free of objection for determining *in vivo* redox potential. But from *in vitro* studies and from the use of the electrode and dye methods, approximations have been made which are, in all probability, very close to the true values.

Cellular Use of Energy

The free energy that results from intracellular oxidation-reduction reactions is available for work, as has already been emphasized. But it should be understood that this free energy cannot be liberated in a large quantity and then somehow stored as a reservoir for future needs. The greatest efficiency results when just sufficient energy is freed to do a specific task, such as synthesizing another compound. As shall be outlined in the next chapter, intracellular metabolism involves reactions that proceed in such a manner that the energy needs of one are supplied by another, which in turn has been supplied by still another. In other words, metabolic processes proceed in a step by step fashion of many interdependent reactions.

SUMMARY

The term metabolism includes all **anabolic** processes by which nutrients are converted into complex tissue elements; the conversion of complex substances to simpler forms is called **catabolism.** These reactions are controlled by biological catalysts termed **enzymes.** The **substrate** is the substance or reaction upon which the enzyme works. All enzymes are thought to be proteins. Enzyme action is influenced by temperature, pH, substrate and enzyme concentration, and the

presence of other chemical substances. Enzyme activity is usually restricted to a specific substrate. Enzymes are thought to act by: 1) forming an intermediate compound with the substrate substance, or 2) lowering the energy of activation. **Energy of activation** is the minimal amount of energy required for a molecule to take part in a reaction.

The metabolic processes are associated with the utilization of oxygen and the liberation of heat. Cellular respiratory rate is expressed as Q_{O_2}. The respiratory rate varies from one unicellular organism to another, and from tissue to tissue. Heat is expressed in **calories.** The large Calorie is equal to 1000 small calories. Oxidation is the ultimate source of energy for all cellular work. The end product is carbon dioxide. The ratio of carbon dioxide produced to oxygen utilized is termed the **respiratory quotient** ($R.Q.$). When the three types of foodstuffs are oxidized, the $R.Q.$ is about 0.82. Heat production of the organism may be determined directly, or indirectly from a determination of the quantity of oxygen consumed.

Oxidation is defined as the loss of an electron or a hydrogen atom. **Reduction** is the acquisition of an electron or a hydrogen atom. The relative ease with which these losses or transfers are made is indicated by the potential of the system, called the **redox potential,** expressed in volts. Substances arranged according to their redox potential form the **electromotive series.** This series indicates in any combination whether a substance will act as an oxidizing or a reducing agent. Any substance or system is capable of oxidizing all systems of lower potential and reducing those of higher potential. Redox potentials are useful in determining the free energy of a reaction. Such energy is generally liberated in small quantities which can be used for other specific intracellular reactions.

Problems

1. Define the following terms:

   a. Metabolism   d. Calorie (large)
   b. Enzyme   e. Respiratory quotient
   c. Energy of activation   f. Redox potential

2. Indicate by the use of a graph the influence of temperature and acidity on enzyme action.

3. In order to oxidize carbohydrate *in vitro* considerable heat must be used. Yet intracellularly carbohydrate is quickly oxidized. Explain.
4. How does knowledge of the *R.Q.* indicate the type of food being utilized?
5. Select two systems from Table 4.4. Indicate which will be the oxidant and which the reductant if they were to interact. Next, calculate the free energy of such a reaction.

Additional Reading

Bray, H. G., and White, K., "Kinetics and Thermodynamics in Biochemistry," Academic Press, Inc., New York, N. Y., 1957.

Davson, H., "A Textbook of General Physiology," Little, Brown and Company, Boston, Mass., 1959.

Laidler, K. J., "Introduction to the Chemistry of Enzymes," McGraw-Hill Book Company, New York, N. Y., 1954.

White, A., Handler, P., Smith, E. L., and Stetten, Jr., D., "Principles of Biochemistry," 2nd ed., McGraw-Hill Book Company, New York, N. Y., 1959.

Wolstenholme, G. E. W., and O'Connor, C. M., eds., "Ciba Foundation Symposium on the Regulation of Cell Metabolism," Little, Brown and Company, Boston, Mass., 1959.

METABOLISM: BASIC REACTIONS

THE BASIC PRINCIPLES of metabolism by which the cells are able to provide energy for their own needs and those of the total organism have been outlined in the previous chapter. The specific reactions by which these energy transformations are carried out will be discussed in this chapter.

The basic foodstuffs are broken down to simpler substances, such as monosaccharides, amino acids, glycerol, and fatty acids. The reactions responsible for these transformations are **exergonic.** An exergonic reaction is one in which the end products possess less free energy than the starting materials. In other words, energy is liberated by an exergonic reaction. In contradistinction, an **endergonic** reaction is one which requires energy to proceed. The end products of endergonic reactions, accordingly, possess more energy than the starting materials. It is important to realize, that within the cell the energy required for endergonic reactions is usually supplied by exergonic reactions.

In most multicellular animals, the first step of metabolism, namely the breakdown of foodstuffs, generally occurs in the digestive tract. The energy liberated by the chemical reactions involved is small, too small to be useful to the cell. It is simply dissipated in the form of heat.

The end products of digestion undergo further transformation within the cell with the liberation of more energy and the production of three substances: 1) acetyl coenzyme A, 2) alpha-ketoglutaric acid,

and 3) oxaloacetic acid. After these three compounds have been produced, they generally enter into the tricarboxylic acid cycle which may be considered to be the final common metabolic pathway. The ultimate end products are water and carbon dioxide.

This degradation of foodstuffs occurs in a series of complex, interlocking, mutually dependent oxidation-reduction reactions in which energy is supplied not only for the reactions to proceed but also for the synthesis of compounds with high energy phosphate bonds. These compounds represent powerhouses, so to speak, from which energy is used for various cellular activities including: 1) metabolic processes, 2) contraction, 3) active transport of substances in and out of the cell, and 4) impulse propagation.

PHASE I

As has just been indicated, in multicellular animals, there are generally three steps, or phases, involved in the series of reactions by which complex foodstuffs are utilized by the organism. Phase I is referred to as **digestion**. Digestion may be simply defined as the act or process of converting food into assimilable form—into products, that is, which the individual cells can use for their metabolic needs.

In phase I of metabolism the complex molecules of carbohydrate, fat, and protein are split. The basic reaction involves the addition of water and therefore is termed **hydrolysis**. Enzymes which participate in these reactions are called hydrolases. There are specific hydrolases for each substrate.

Hydrolysis of Carbohydrate

Carbohydrates appear in many forms, ranging from the simple monosaccharides, such as glucose, on up to the complex polysaccharides, such as starch. The hydrolysis of starch and other polysaccharides is catalyzed by **amylase**. In man there is amylase in saliva and also in pancreatic juice, both of which enhance the splitting of the polysaccharides to **oligosaccharides**. The combining form, oligo-, means "few." Combined with saccharide it means few saccharides and therefore refers to molecules that are not as complex as starch and glycogen. In general, a carbohydrate with 2 to 10 monosaccharide units is termed an oligosaccharide.

Ultimately, hydrolysis proceeds, under the influence of the amylases, until maltose, sucrose, and lactose result. These disaccharides are then further hydrolyzed to monosaccharides under the influence of the enzymes **maltase, sucrase,** and **lactase.**

The end products of carbohydrate digestion are glucose, galactose, and fructose. Glucose predominates by far.

Hydrolysis of Lipid

The lipids, which must be hydrolyzed, include the neutral fats, phosphatides, steroids, and the fat-soluble vitamins, A, D, E, and K. In the diet, the neutral fats are in greatest quantity. Neutral fat is hydrolyzed under the influence of **lipases,** to glycerol and fatty acids. In Fig. 5.1, it can be seen that in this hydrolysis three fatty acids and one molecule of glycerol result.
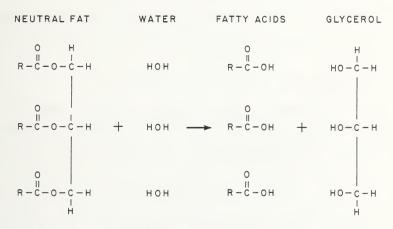
Fig. 5.1.   Hydrolysis of Neutral Fat to Fatty Acids and Glycerol.

The surface tension between fat and water is very high. For this reason hydrolysis of fat, even in the presence of specific enzymes, does not occur readily. In the intestinal tract it is the function of bile to reduce this surface tension and to aid in emulsification so that the fat droplet is broken up into very small drops of fat which present a much greater surface area than does a single relatively large mass. Thus, because more surface is exposed and because the surface tension is reduced, the hydrolysis proceeds effectively.

Hydrolysis of Protein

In man there are specific enzymes in the gastric juice, in the pancreatic juice, and in the intestinal juice which catalyze various stages of the digestion of protein. **Pepsin,** in the gastric juice, initiates hydrolysis. This enzyme, if given optimal conditions and sufficient time, is capable of carrying the hydrolysis all the way to the amino-acid stage, but in the stomach, polypeptides generally result. The polypeptides are then further hydrolyzed in the small intestine by **trypsin** which is secreted in the pancreatic juice. There are also other proteinases and peptidases in the pancreatic juice and also in the intestinal juice which assure that all proteins and polypeptides are hydrolyzed to the end products of protein digestion, the amino acids.

PHASE II

The digestion of the foodstuffs is considered to be phase I. The next step is to degrade further the end products of hydrolysis. This occurs exclusively within the cell.

Glucose Phosphorylation

Glucose oxidation liberates a tremendous burst of energy with the production of water and carbon dioxide. To be sure, this energy is ultimately released, and water and carbon dioxide are produced, but this occurs through a series of highly complicated interlocking steps. One of the first steps for many of the reactions involving glucose is phosphorylation, that is, the formation of a compound containing glucose and phosphate. Adenosine triphosphate (ATP), a compound with high energy phosphate bonds, generally enters into this reaction as follows:

$$\text{ATP} + \text{Glucose} \xrightarrow{\text{hexokinase}} \text{Glucose 6-phosphate} + \text{ADP}$$

ADP stands for adenosine diphosphate, a substance which differs from ATP in that it has one less high energy phosphate bond. It will be noted that the ATP + glucose reaction is catalyzed by the enzyme **hexokinase.** This type of enzyme is also known as **glucokinase.**

Glucose moves through the cell wall more readily than does glucose 6-phosphate. Thus, not only does phosphorylation serve to prepare

glucose for further metabolic processes, but it also prevents it from diffusing out of the cell.

Glycolysis

Glucose 6-phosphate can be directly hydrolyzed by some cells so as to liberate the glucose molecule once again. This certainly occurs in the liver, for example, and constitutes one of the mechanisms for maintaining the blood sugar constant. The hydrolysis of glucose 6-phosphate is catalyzed by a specific **phosphatase.**

Another pathway for the release of energy contained in the glucose molecule is depicted in Fig. 5.2. The end product of this sequence is pyruvic acid. It should be noted that in this entire series of complex transformations, no oxygen is used. It is for this reason that the sequence is termed **anaerobic.**

When sufficient oxygen is present, the pyruvic acid is converted to **acetyl coenzyme A** which is a major end product of phase II. Pyruvic

Glucose

↓

Glucose 6-phosphate

↓

Fructose 6-phosphate

↓

Fructose 1,6-diphosphate

↓

Glyceraldehyde 3-phosphate

↓

2 1,3-Diphosphoglyceric acid

↓

2 3-Phosphoglyceric acid

↓

2 2-Phosphoglyceric acid

↓

2 Phosphopyruvic acid

↓

2 Pyruvic acid
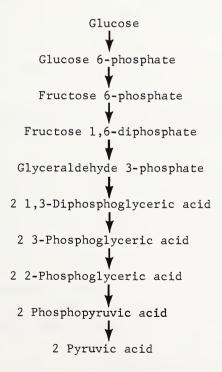
Fig. 5.2. Anaerobic Glycolysis of Glucose.

acid, instead of being converted to coenzyme A, can be reduced to form lactic acid which can then undergo oxidation to reform pyruvic acid. In other words, if plenty of oxygen is present, the pyruvic acid continues into the metabolic mill for the final transformation to water and carbon dioxide. But if sufficient oxygen is not available, then lactic acid is formed and accumulates until oxygen is available to reconvert it to pyruvic acid and then on to other products.

In this connection the **Pasteur effect** should be mentioned. Pasteur in 1861 observed, and his observation has since been amply confirmed, that the rate of glycolysis is lower under aerobic than under anaerobic conditions. It is the inhibition of glycolysis by oxygen that is called the Pasteur effect. At first thought, such inhibition may seem to be undesirable, but actually it has a very beneficial effect. This follows from the fact that less than 10 percent of the energy of aerobic respiration can be obtained from anaerobic respiration. Thus, when adequate oxygen is available, the energy needs of the cells are met so much more efficiently than they are in the absence of oxygen that far less glucose is used. To put it another way, glycolysis is inhibited, glucose is conserved, and still the oxygen needs are satisfied. But exactly how the Pasteur effect operates is not yet understood.

If glucose is completely oxidized, 4.1 Calories per gram are liberated. But if no oxygen is available, the anaerobic transformation to lactic acid produces less than 10 percent of that potential. The major energy yield, then, comes about when lactic acid is oxidized to pyruvic acid and on to water and carbon dioxide. Generally all of the lactic acid is not used in this way. In muscle, at any rate, only about 25 percent of the lactic acid is completely oxidized. The energy so liberated is then utilized to reverse the reactions shown in Fig. 5.2. As a result, glucose 6-phosphate is reformed and, as will be discussed later, this substance is then transformed into glycogen.

It will be noted that fructose 6-phosphate results from the transformation of glucose 6-phosphate. Fructose 6-phosphate can also arise directly from fructose which is also one of the end products of carbohydrate digestion. This reaction is:

$$\text{ATP} + \text{Fructose} \xrightarrow{\text{fructokinase}} \text{Fructose 6-phosphate} + \text{ADP}$$

Finally, it should be appreciated that each step in the series of reactions leading to the formation of pyruvic acid from glucose or fructose requires a specific enzyme.

Fatty Acid Oxidation

The end products of neutral fat digestion are fatty acids and glycerol. The glycerol is used mostly for the biosynthesis of triglycerides and other lipid products. The fatty acids are oxidized with the liberation of considerable energy. It will be recalled that fat yields 9.3 Calories per gram when completely oxidized.

Oxidation of fatty acids occurs at the so-called beta-carbon atom (Fig. 5.3). As a result of this oxidation, a molecule of acetic acid is

Fig. 5.3.   Beta-Carbon Oxidation of Fatty Acid.

split off. Continued oxidation of the fatty acid shortens the molecule by two carbon atoms each time until the entire fatty acid molecule is oxidized.

The oxidation of fatty acid is not truly as simple as Fig. 5.3 indicates. It is now believed to require at least four successive reactions involving specific enzymes with the final products being not simply acetic acid, but rather acetyl coenzyme A. This means that the end product of monosaccharides as well as fatty acids in phase II is acetyl coenzyme A.

Amino Acid Deamination

The first step in the oxidation of amino acids involves the removal of the amino radical. This is termed deamination. The reaction is:

In mammals this reaction occurs mostly in the liver, but it also takes place to a limited extent in the kidney and intestinal mucosa.

Keto Acid Transformation

The deamination of amino acids gives rise, as shown above, to ammonia and a keto acid. The keto acids then undergo various transformations. Leucine, tyrosine, and phenylalanine can be converted to acetyl coenzyme A. Glutamic acid, histidine, proline, and arginine are changed to **alpha-ketoglutaric acid.** Aspartic acid, tyrosine, and phenylalanine may undergo a series of reactions leading to **oxaloacetic acid.**

Thus, in the phase II of the metabolic process the end products of digestion are converted to acetyl coenzyme A, alpha-ketoglutaric acid, and oxaloacetic acid. Three products are now ready to enter into the final common metabolic pathway.

PHASE III

The tricarboxylic acid cycle, which constitutes phase III, is also known as the **citric acid cycle,** or the **Krebs cycle** after the investigator who was responsible for elucidating so many of its intricate interrelationships. In this cycle, the closely related end products of phase II may be completely oxidized to CO_2 and H_2O. In Fig. 5.4 it is seen that acetyl coenzyme A reacts with oxaloacetate to form citric acid and to liberate the coenzyme. The citric acid then undergoes several transformations until oxaloacetic acid is once again formed. These are simply the major steps. In each step there are several interlocking reactions all of which are catalyzed by specific enzymes.

In this cycle only the "acetyl" part of acetyl coenzyme A is used up. So long as this component keeps entering the cycle, the reactions will continue. In one revolution 2 molecules of oxygen are consumed and 2 molecules of carbon dioxide are produced.

Although the cycle will continue to operate indefinitely so long as oxygen is available and acetyl coenzyme A is fed into it, the other products of phase II metabolism, namely alpha-ketoglutaric acid and oxaloacetic acid, may enter into the cycle as indicated in Fig. 5.4. Because all three end products may take part, the tricarboxylic acid cycle is truly the final common metabolic pathway.

Respiratory Enzymes

In so far as is known, each respiratory enzyme functions in conjunction with a coenzyme. By definition, a coenzyme is simply a substance associated with and activating an enzyme.

The number of specific enzymes taking part in the degradation of the foodstuffs through the three phases of metabolism is truly great. No attempt will be made to identify each one. They may be considered as belonging to various categories.

1. *Hydrolases.* As has already been pointed out, the hydrolases catalyze the reaction in which water splits the complex organic molecules into simpler compounds.

2. *Oxidases.* Broadly speaking, an oxidase is an enzyme that promotes an oxidation reaction. More specifically, after the oxidation
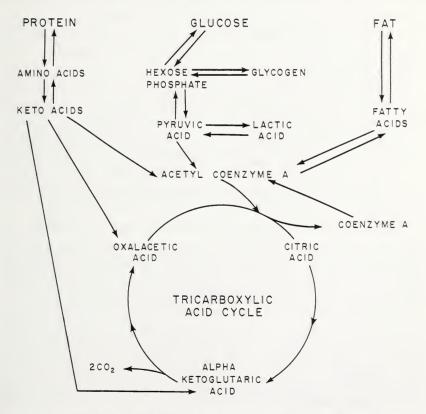
Fig. 5.4. Interconversions of Protein, Glucose and Fat.

reaction has been completed as a result of the presence of an oxidase, the reduced O_2 appears in the form of peroxide or water thus:

$$H_2 + O_2 \longrightarrow H_2O_2$$
or
$$4H+ + O_2 \longrightarrow 2H_2O$$

Cytochrome oxidase is an example of this group. As the term suggests, cytochrome is a cell pigment. It is an iron-containing protein. Actually, there are several compounds classified as cytochromes which are chemically very similar and which have in common the ability to be reversibly oxidized with ease. They are differentiated by spectroscopic means. In the reduced state they exhibit characteristic and sharp absorption bands. Letters are used to identify the cytochromes. Originally, there were thought to be but three, known as cytochrome a, cytochrome b, and cytochrome c. It is now known that there are others and so they are designated as a_2 or b_2, or perhaps b_3, etc., depending upon which basic type the particular one resembles most. Of them all, cytochrome c is the most abundant.

Only one of the cytochromes is known as cytochrome oxidase. This compound is oxidized by molecular oxygen and then, it in turn oxidizes the other cytochromes. The compound identified according to its absorption band as cytochrome a_3 is believed to be cytochrome oxidase. In the process by which the cytochromes are oxidized there is a passage of electrons from the substrate to the oxygen, and the iron in each cytochrome molecule is oxidized from the ferrous to the ferric form.

Oxidation-reduction interrelationships are commonly written as in Fig. 5.5. It is seen that molecular oxygen reacts with ferrous iron in cytochrome oxidase. As a result the molecular oxygen is reduced to oxygen ion, and ferrous iron is oxidized to the ferric form. A respiratory chain can then be built by additional units (Fig. 5.6).

The cytochromes are found in most cells, but they are particularly concentrated in muscle cells which suggests that cytochromes are essential for reactions which must proceed rapidly and with a sudden burst of energy.

In addition to the cytochromes, there are oxidases which are copper-containing proteins.

3. *Dehydrogenases.* If the term oxidase is used in a more limited sense than it was used above, it would be defined as a substance that facilitates electron transfer. In this sense, then, a dehydrogenase is concerned with hydrogen-ion transfer. As the term, dehydrogenase, indicates these are enzymes which act upon various steps in the process of dehydrogenation. Most organic compounds are oxidized in this way. In most respiratory chains oxidases and dehydrogenases are both essential to the interlocking reactions.

Diphosphopyridine nucleotide (DPN), coenzyme I, and triphosphopyridine nucleotide (TPN), coenzyme II, are synthesized by many

$$2 \text{ CYTOCHROME OXIDASE } Fe^{++} + \tfrac{1}{2} O_2 \longrightarrow$$

$$2 \text{ CYTOCHROME OXIDASE } Fe^{+++} + O^=$$

$$O^= + 2H^+ \longrightarrow H_2O$$
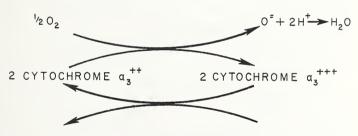
Fig. 5.5. Oxidation-Reduction Reactions. Above, the older way of expressing the reaction. Below, the newer form showing the interrelationships.

cells. These pyridine nucleotides are coenzymes which function quite commonly with the flavin-containing enzymes, the so-called **flavoproteins.** Flavin adenine dinucleotide seems to be the essential component of these enzymes and therefore the abbreviation FAD is often used.

The cytochromes are not believed to act directly on the substrate. This is, apparently, the role of the dehydrogenases. In a typical respiratory chain (Fig. 5.6) the substrate is oxidized. At the same time DPN is reduced. DPN now reacts with flavoprotein to be oxidized

while the flavoprotein is reduced. In the next step, 2 molecules of cytochrome b are reduced with the liberation of 2 hydrogen ions. The electrons then are handed down the cytochrome chain until finally molecular oxygen is reduced to oxygen ion which reacts with the 2 hydrogen ions to form water. In this chain, energy is liberated.

$\frac{1}{2} O_2$ $O^= + 2H^+ \longrightarrow H_2O$

2 CYTOCHROME a_3^{++} 2 CYTOCHROME a_3^{+++}

2 CYTOCHROME a^{++} 2 CYTOCHROME a^{++}

2 CYTOCHROME c^{++} 2 CYTOCHROME c^{+++}

2 CYTOCHROME b^{+++} 2 CYTOCHROME $b^{++} + 2H^+$

FLAVOPROTEIN   FLAVOPROTEIN
2e.2H

DPN   DPN $+$ 2e.2H

SUBSTRATE   OXIDIZED
2e.2H   SUBSTRATE
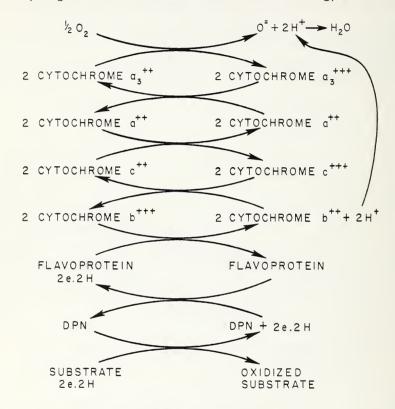
Fig. 5.6. A Typical Respiratory Chain.

4. *Transferases.* Enzymes which catalyze the transfer of a radical from one compound to another are known as transferring enzymes, or transferases. Radicals capable of being transferred include methyl, phosphate, and amide groups.

5. *Isomerases.* These are enzymes which catalyze internal changes in a molecule. For example, the change from glucose 6-phosphate to fructose 6-phosphate, or glucose 1-phosphate to fructose 6-phosphate.

6. *Decarboxylases.* Under the influence of the decarboxylases, car-

bon dioxide can be removed from a molecule of carboxylic acid without oxidation.

Phosphorylation

Energy is released at several points along a respiratory chain. This energy seems to be used primarily for the transformation of adenosine diphosphate (ADP) to adenosine triphosphate (ATP). It is thought that the oxidative phosphorylation of ADP occurs in at least three places in the respiratory chain: 1) in the electron transport from DPN to flavoprotein, 2) in the electron transport from cytochrome b to cytochrome c, and 3) in the electron transport from cytochrome c to cytochrome a_3. In the tricarboxylic acid cycle there are several respiratory chains.

It has been calculated that in one cycle there are 12 moles of ATP formed. The synthesis of one mole of ATP from ADP requires approximately 7 Calories. This means that in one cycle some 84 Calories are utilized to synthesize ATP from ADP. The full energy yield of a mole of acetic acid is over 200 Calories. In other words, about 60 percent of the energy is utilized for chemical interreactions in the cycle, for other cellular purposes, and is dissipated as heat; whereas 40 percent of the total energy potential is conserved in the form of ATP. The cell, with a supply of ATP, now has a compound that is rich in energy and which, by virtue of its phosphate bonds, can quickly liberate a large burst of energy for all cellular energy needs.

In brief, by virtue of the many interreactions the energy requirements of the cell are met. In addition, except when the cell is involved in exhaustive activity, an energy pool consisting largely of organic phosphates, especially ATP, results. This pool, or reservoir, is ever available to supply sudden cellular energy demands.

BIOSYNTHESIS

Thus far, with the exception of the synthesis of ATP from ADP, catabolism has been the primary consideration. The metabolic fate of the basic foodstuffs has been traced to their ultimate end, complete oxidation to carbon dioxide and water with the liberation of energy. Another cellular function is the formation of complex compounds from simpler ones, that is, biosynthesis.

Photosynthesis

Light is a source of energy that can be used by some cells for biosynthesis. Because light is involved, this type of biosynthesis is termed photosynthesis.

The over-all reaction involves carbon dioxide and water with the production of carbohydrate and oxygen, thus:

$$CO_2 + H_2O \xrightarrow{\text{light}} CH_2O + O_2$$

Examination of this reaction gives one the impression that the carbon dioxide molecule has been split. But by the use of an isotope of oxygen, O^{18}, it has been shown that the carbon dioxide is reduced and it is the water that is split, thus:

$$CO_2 + 2H_2O^{18} \xrightarrow{\text{light}} CH_2O + H_2O + O_2^{18}$$

It can be seen that the 2 molecules of water are completely split, liberating one molecule of oxygen. Two hydrogens from the split water molecule combine with oxygen from CO_2, while the other two combine to form the basic carbohydrate CH_2O. This, then, is an oxidation-reduction reaction in which the carbon dioxide is reduced and the water oxidized.

Water may be oxidized by other substances in the absence of CO_2. The substance must be an electron acceptor. Thus, cytochrome may play a role in the conversion of ferric to ferrous iron. Or DPN may be involved with the production of reduced DPN.

It has been shown that photosynthesis occurs in two steps:

1. *Light reaction.* Under the influence of light in the presence of an electron acceptor, water is split with the liberation of oxygen:

$$2DPN + 2H_2O \xrightarrow{\text{light}} 2DPNH_2 + O_2$$

Since this reaction, in order to proceed, requires light, it is termed the light reaction. There is still some disagreement as to the exact products formed in the light reaction. $DPNH_2$ is no doubt formed, but the photosynthetic phosphorylation of ADP has also been demonstrated. Thus, ATP could also be a product of the light reaction.

2. *Dark reaction.* The electron acceptor then reduces carbon dioxide:

$$2DPNH_2 + CO_2 \longrightarrow 2DPN + H_2O + CH_2O$$

This reaction proceeds without the radiant energy of light and therefore is termed the dark reaction.

A green pigment found in many plant cells, namely chlorophyll, seems to be essential for the light reaction. It is possible that other pigments may serve the same purpose. The pigment does not necessarily undergo a chemical change. Its role seems to be to absorb light. The light then is made available, in some manner not clearly understood, to provide energy for the light reaction. It is believed that the other cell pigments, after absorbing light, transfer it to chlorophyll before it can be used for the light reaction.

The structure of chlorophyll is shown in Fig. 5.7. There are two types of chlorophyll, chlorophyll a and chlorophyll b. They are thought to differ in that chlorophyll b has one CH_3 group replaced by a CHO group.

The rate at which photosynthesis proceeds depends upon: 1) the intensity and wavelength of the light, 2) the concentration of carbon
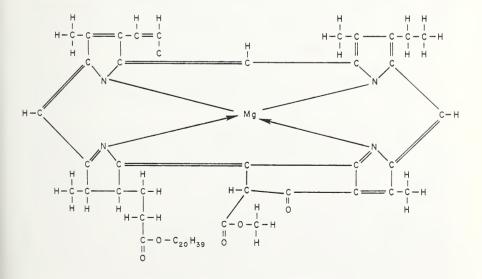
Fig. 5.7.   The Structure of Chlorophyll.

dioxide, 3) the amount and type of pigment in the cell, and 4) the temperature.

An idea of the amount of energy required for photosynthesis can be had by a consideration of the fact that the complete oxidation of 1 gram of glucose produces 4.1 Calories. Clearly, then, the photosynthesis of a gram of glucose requires at least that much energy. This energy comes from the sunlight. The importance of photosynthesis in the total life processes can thus be appreciated.

Carbohydrate Biosynthesis

Photosynthesis is one example of carbohydrate biosynthesis. But this only occurs, it is believed, in plant cells—in cells that possess light-absorbing pigments such as chlorophyll. The animal cell depends upon the external environment for its fundamental carbohydrate supply. Such cells, as has already been discussed, can utilize carbohydrate for energy needs. They can also convert simple carbohydrates into more complex ones.

Plant cells synthesize starch from the simple sugars: animal cells form glycogen. Smaller molecules, the oligosaccharides, also are formed. The formation of glycogen is termed **glycogenesis.**

The fundamental reaction in the biosynthesis of glycogen is the conversion of glucose to glucose 6-phosphate:

$$\text{Glucose} + \text{ATP} \longrightarrow \text{ADP} + \text{Glucose 6-phosphate}$$

It is to be observed that the energy required for this reaction is provided by ATP which gives up one high energy phosphate bond in the transformation to ADP.

The next step is a relatively simple internal change of glucose 6-phosphate to glucose 1-phosphate. Many units of glucose 1-phosphate then are combined into the glycogen molecule after giving up their phosphate. This, of course, is a major transformation about which very little is known. There is some evidence that there is still another intermediary. Some authorities believe that glucose 1-phosphate is converted to uridine diphosphate-glucose and then this product is utilized to synthesize glycogen. These possibilities are depicted in Fig. 5.8. Each step is catalyzed by specific enzyme systems.

Glycogen serves at least two purposes: 1) it represents a reservoir of glucose and thus functions to regulate the blood sugar by supply-

ing glucose when ingestion is too low and by removing glucose from the circulation when blood sugar rises and 2) it is a source of energy. It plays a particularly important role in this respect in the chemistry of muscular contraction.
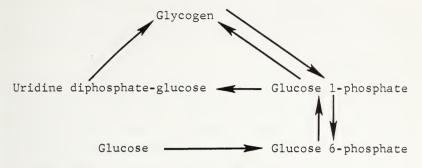
Fig. 5.8. Glycogen Synthesis.

Lipid Biosynthesis

The digestion of neutral fat supplies fatty acids which can be utilized for lipid biosynthesis. The cell is also capable of synthesizing fatty acids from carbohydrate and protein sources. Even though the cell can convert glucose into fatty acids, the reverse reaction does not seem to occur. The reason for this is seen in Fig. 5.4 in which the metabolic interconversions are shown. Pyruvate forms acetyl coenzyme A, but this reaction is not reversible. Thus glucose can go through pyruvate to acetyl coenzyme A to fatty acid, but the fatty acids can only go to acetyl coenzyme A, but not to pyruvate and therefore on to glucose.

The fatty acids, whether resulting from the digestion of fat, from the conversion of glucose, or from deaminated amino acid residues, may then be conjugated with glycerol to form neutral fat. It is in this form that it is deposited in various tissues, but especially in adipose tissue.

Fatty acids are synthesized into products other than neutral fat. These include the **phospholipids,** of which lecithin, cephalin, and sphingomyelin are the most important. Phospholipids contain fatty acid, phosphoric acid and a nitrogenous base. These substances serve to transport fatty acids, they act as insulators around nerve fibers, they take part in the blood-clotting process and they act as phosphate

donors. They also make up essential physical structures of the cell, such as the cell and nuclear membranes.

Cholesterol is another important lipid. It is supplied in the diet, but it can also be formed by biosynthesis. It has been shown that all of the carbon atoms of synthesized cholesterol are derived from acetyl coenzyme A. Cholesterol is also used for essential cell structures. In addition it seems to be a precursor of some of the steroid hormones.

Protein Biosynthesis

The amino acids that result from the digestion of protein can be utilized by the cell for several purposes. One of these is the biosynthesis of protein. It has been shown that protein is constantly being degraded to amino acid and, at the same time, protein is synthesized from amino acids. There is thus a constant turnover of protein. The exact mechanism for protein synthesis is not known. It has been shown that the energy requirements for the formation of a peptide linkage are quite high. Certain peptide bonds require over 3 Calories for their formation. When one takes into consideration the great number of such bonds in a protein molecule, it can be appreciated that considerable energy is required for protein synthesis. The remarkable aspect is that despite the complexity of the protein molecule and the great energy demands, most cells are capable of synthesizing protein readily and rapidly. The nucleoproteins, discussed in Chapter 1, are believed to play an essential role in the control of protein synthesis. They are thought to act as autocatalysts, that is they hasten their own formation and, in this way, reproduce themselves.

The biosynthesis of protein is essential for growth, for the production of plasma protein, and for the formation of enzymes and hormones. Since amino acids probably are not formed from carbohydate or lipid sources, it becomes obvious why protein is an essential constituent of the diet.

BIOLUMINESCENCE

It has been noted that plant cells are capable of utilizing radiant energy for biosynthesis. In many cells the reverse reaction occurs, that is, light is emitted by a chemical reaction. The emission of light by a living organism is termed bioluminescence.

Luminescent Organisms

Luminescence has been reported in plants, but it is rare and generally very weak. Luminescence in animal cells is far more common and certainly more spectacular.

Luminescence has been observed in a wide range of organisms extending from bacteria through the vertebrates. Probably the first observations of bioluminescence were made by sailors who noted at night that the sea churned by their ship glowed brilliantly. This is not due to luminescent bacteria but rather to protozoa. *Noctiluca miliaris* is but a single-celled protozoan less than 1 mm in diameter, yet it is so concentrated in sea-water that its glow is readily detected.

There are many other luminescent animals including jellyfish, fishes, and crustaceans, but perhaps best known is the firefly. These organisms have a light-organ that gives a truly brilliant flash of light. The organ is under the control of the nervous system.

Function of Luminescence

Luminescence is thought to serve three functions: 1) sexual attraction, 2) protection by blinding, and 3) attraction of prey. In the firefly the sexual attraction of the luminescence has been shown very clearly. At night the male flashes and the female then flashes in response. Apparently to make certain, the male then flashes again and awaits the response from the female. This sequence is repeated four or five times which seems to suffice for complete assurance, for thereafter they mate. One wonders how the male can tell the responding flash of the female from a flash from another male. This has been shown to involve a time relationship and not any difference in the quality of the flash. It probably explains why the sequence is carried out four or five times. After the male flashes, the female responds in exactly 2 seconds with her own flash. Such a flash could come purely inadvertently from another male. However, when the flash and response is repeated 4 or 5 times and each time the time interval is 2 seconds, practically all possibility of chance flashes from another male is eliminated, or so reasons the statistically minded firefly who then moves in to mate!

In other forms, for example, in the marine worm, it is the female who flashes first to attract the male. In this case the female swims in

a circle, all the while luminescent. The male is attracted, approaches, glows, and then together they continue luminescent while the eggs are fertilized.

For protection, some forms can secrete a material which becomes luminescent upon contact with sea water. Thus, when these organisms are being pursued, they secrete the luminescent material to blind their pursuer and then make their escape.

Finally, it is thought that in some instances luminescence serves no purpose but is merely a by-product of metabolic processes.

The Nature of Bioluminescence

The absolute candle power of luminous organisms is low, ranging from 1/50 to 1/400 candle. However, when the candle power is calculated on the basis of unit of luminous surface, the efficiency is found to be considerable.

The color of the light varies with most luminescence occurring in the orange-yellow-green-blue range. If the spectral energy is plotted (Fig. 5.9), it can be seen that there is a rather narrow spectrum with
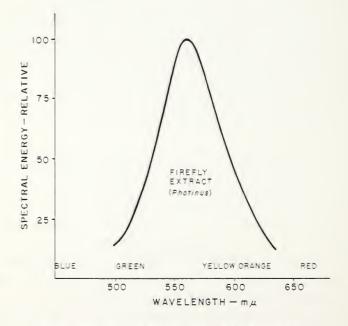
Fig. 5.9. Spectral Energy Curve of Firefly Extract.

a definite peak. In this case, for one of the fireflies the peak is seen to be at about 560 millimicrons which is in the yellow-green range.

Mechanism of Luminescence

Luminescence is thought to occur due to the oxidation of a substance that many years ago was termed **luciferin.** This reaction is catalyzed by the enzyme **luciferase.** Luciferin is apparently stored as granules which must first be broken down before the luminescent reaction can proceed. But the precise mechanism by which this occurs is not known.

It is interesting that the mixture of luciferin from one organism with luciferase from another is without effect, that is, no luminescence results. unless the two organisms are closely related.

It has been postulated that luminescence in the firefly is controlled by regulating the admission of oxygen to the reactive system. Thus, it is believed that the following reactions occur:

$$\text{Inactive Intermediate} \xrightarrow{\text{Luciferase} + \text{ATP} + \text{Luciferin} + \text{Mg}^{++}} \text{Active Intermediate}$$

$$\text{Active Intermediate} \xrightarrow{O_2} \text{Light}$$

In the absence of oxygen the active intermediate, of a still unknown nature, accumulates. When oxygen is admitted, there is a flash of light. But whether luminescence is controlled by regulating oxygen admission or by some other mechanism is not certain. It is thought possible that the nervous control over luminescence is similar to its control of muscle contraction. If so the liberation of ATP is the key mechanism.

SUMMARY

A reaction that liberates energy is said to be **exergonic;** one that requires energy to proceed is **endergonic.** Phase I of metabolism, the hydrolysis of the basic foodstuffs during digestion involves exergonic reactions. The end products of carbohydrate digestion are the **monosaccharides;** of lipid digestion, **glycerol** and **fatty acids;** and of protein digestion, the **amino acids.**

In phase II, the monosaccharides, fatty acids, glycerol, and some

of the amino acids are degraded to **acetyl coenzyme A.** Other amino acids form **alpha-ketoglutaric acid** and **oxaloacetic acid.**

In phase III, the end products of phase II enter into the **tricar-boxylic acid cycle.** This is a cycle of interrelated, or coupled reactions. All of the constituents which take part in the cycle are regenerated with the exception of acetyl coenzyme A. So long as oxygen is available and acetyl coenzyme A is abundant, the cycle will continue to operate indefinitely with the production of H_2O and CO_2, and with the liberation of considerable energy. For these oxidation-reduction interactions many specific enzymes are required. These include the **oxidases** and **dehydrogenases.** The **cytochromes** are important oxidases. The **flavoproteins** are dehydrogenases which function in conjunction with coenzyme I, **diphosphopyridine nucleotide (DPN)**, and coenzyme II, **triphosphopyridine nucleotide (TPN)**.

The energy released during the tricarboxylic acid cycle is used for the transformation of **adenosine diphosphate (ADP)** to **adenosine triphosphate (ATP)**. In one cycle there is enough energy liberated for the synthesis of 12 moles of ATP representing some 84 Calories of energy. ATP is an energy-rich compound which can liberate a large burst of energy for all cellular needs.

Biosynthesis is the formation of complex molecules from simpler ones by the cell. Plant cells are capable of **photosynthesis,** a process which uses the radiant energy of light for the formation of carbohydrate from CO_2 and H_2O. The monosaccharides can be converted by intracellular processes to oligosaccharides and glycogen. Fatty acids can be conjugated with glycerol to form neutral fat. They can also be combined with phosphoric acid and a nitrogenous base to form the **phospholipids.** Another product of lipid biosynthesis is **cholesterol.** Amino acids are used by the cell for the synthesis of polypeptides and protein.

Many organisms can give rise to **bioluminescence.** This serves for: 1) sexual attraction, 2) protection, and 3) attraction of prey. The color of the light is generally in the orange-yellow-green-blue range of the spectrum. Luminescence involves the oxidation of **luciferin** catalyzed by **luciferase.** Oxygen and ATP are essential for this reaction and may be involved in the nervous control of luminescence.

Problems

1. What functions does the phosphorylation of glucose serve?
2. Construct a respiratory chain, identify each substance that is oxidized and each that is reduced, and explain how the substrate is oxidized with the formation of water.
3. Why is the tricarboxylic acid cycle called the final common metabolic pathway?
4. Explain the roles that the oxidases and dehydrogenases play in metabolic processes.
5. Starting with a mole of glucose, outline the energy transformations that occur before complete degradation to CO_2 and H_2O occurs.
6. Explain the difference between the light and dark reactions in photosynthesis.

Additional Reading

Brody, S., "Bioenergetics and Growth," Reinhold Publishing Corporation, New York, N. Y., 1945.

Davson, H., "A Textbook of General Physiology," Little, Brown and Company, Boston, Mass., 1959.

Greenberg, D. M., ed., "Chemical Pathways of Metabolism," vol. II, Academic Press, Inc., New York, N. Y., 1954.

Johnson, F. H., ed., "The Luminescence of Biological Systems," Am. Assoc. Advance. Sci., Washington, D. C., 1955.

Krebs, H. A. and Kornberg, H. L., "Energy Transformations in Living Matter," Springer-Verlag, Berlin, 1957.

Strominger, J. L., "Mononucleotide Acid Anhydrides and Related Compounds as Intermediates in Metabolic Reactions," *Physiol. Rev.,* 40, 55-111 (1960).

MOVEMENT

ONE OF THE MOST fundamental and characteristic properties of cells, especially animal cells, is the ability to move. This movement may be little more than the movement of protoplasm within the cell, a phenomenon termed **protoplasmic streaming.** There is little or no movement of the cell itself. About the simplest form of actual cell locomotion is exhibited by a single cell such as the ameba. A more advanced type of movement is brought about by cells that possess cilia or flagella. And finally, there are the animal forms which have muscles—tissue with the primary purpose of movement of the organism.

Very little is known concerning the mechanism of protoplasmic streaming, ameboid, ciliary, or flagellar movement. Thus, little more than a description and some of the factors that influence these types of movement can be given. In contrast there is a tremendous body of knowledge concerning all aspects of muscular movement and, therefore, the greater part of this chapter will be devoted to that subject.

PROTOPLASMIC STREAMING

Protoplasmic movement most often takes the form of a rotary motion. Under the microscope, protoplasm is often seen to move around a vacuole. Because of this rather characteristic rotary movement, the phenomenon is also known as **cyclosis.**

Function of Streaming

Just what function protoplasmic streaming carries out is not definitely known. It seems most probable that it serves to transport

essential substances throughout the cell. In this light it would be an important adjunct to diffusion and active transport. In contrast to this view is the conclusion of some authorities that protoplasmic streaming does not normally occur, thus has no physiological importance, and is only observed when the cell has been injured.

Factors that Influence Protoplasmic Streaming

Against the argument that protoplasmic streaming is but a reaction to injury are the observations of the many types of stimuli that alter streaming, often in a very quantitative way. These stimuli include:

1. *Temperature.* If the rate of protoplasmic streaming is plotted against the environmental temperature, a straight line curve results, up to a point (Fig. 6.1). Whether this is simply due to a change in the viscosity of the protoplasm in response to temperature, or to some other factors, is not known.

2. *Light.* Streaming may be initiated as well as altered by light of certain wavelengths, intensity, and duration. The initiation of protoplasmic streaming by light is termed **photodinesis.**

3. *Respiratory Gases.* Protoplasmic streaming can continue, for a variable period of time, in the absence of oxygen. Ultimately it stops
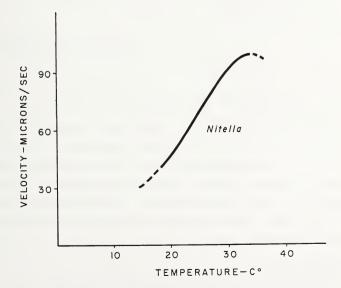
Fig. 6.1. The Influence of Temperature on Protoplasmic Streaming.

if oxygen is not provided. This observation strongly supports the contention that streaming is a metabolic phenomenon rather than simply a change in viscosity, or a response to injury. The influence of carbon dioxide has not, apparently, been adequately studied. High concentrations of carbon dioxide stop streaming, but what influence more physiological levels of the gas have is not known.
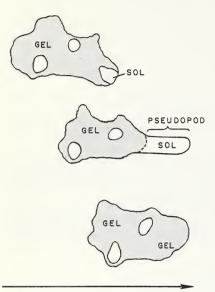
Fig. 6.2. Ameboid Movement. In one area of the cell there appears to be a change from the gel to the sol phase. This permits a pseudopod to form and then the remainder of the cell flows into it, thus causing the entire cell to move.

4. *Chemicals.* A wide variety of chemical agents have been used in this study. Clearly, streaming can be influenced in this way. The initiation of protoplasmic streaming by chemical agents is called **chemodinesis.**

5. *Mechanical and Electrical Stimuli.* Electrical stimuli generally cause stoppage of protoplasmic streaming, but mechanical stimuli are apparently without influence until the intensity becomes so great as to be highly unphysiological.

AMEBOID MOVEMENT

In protoplasmic streaming, the cell as a whole does not move. However, the protoplasm of the ameba also exhibits a flowing action, but as a result, the entire organism moves. Although this type of movement is so very characteristic of the ameba and thus is called ameboid movement, single cell organisms, other than the ameba, also move in this way. In mammals the white cells that possess phagocytic properties exhibit this type of movement.

Means of Locomotion

The mechanism of ameboid movement is not clear. This single cell organism has no fixed shape. As one watches, the shape is seen to

change and, apparently by virtue of such changes, the position of the ameba is altered. It appears as though a projection of the cell is sent out in a particular direction, a so-called **pseudopod** ("false foot"), and then the remainder of the protoplasm simply flows into it (Fig. 6.2).

Ameboid movement may result from progressive solation and gelation. The pseudopod, according to this view, develops because the cortical gel at one point changes to a sol. There is now no support at that point, and, therefore, it balloons out. After this occurs, gelation takes place giving the pseudopod shape and rigidity. Whether the remainder of the cell simply flows into it, or whether there is an actual contraction that pulls the cell forward, is not settled.

The rate of movement varies with the type of organism and the environmental conditions, but it seems to average around 0.2 mm per minute.

Factors that Influence Ameboid Movement

The response of the ameba to various stimuli is very similar to the influence these factors have on protoplasmic streaming. Thus the rate of movement increases with temperature up to about 35°C and then falls off. Likewise activity can be stopped by the lack of oxygen, or excessive carbon dioxide, or strong mechanical and electrical stimuli. It would appear as though the same basic protoplasmic mechanisms are responsible for both types of movement.

CILIARY AND FLAGELLAR MOVEMENT

From a physiological standpoint the primary difference between cilia and flagella is length. Cilia are short, flagella are long. These cellular appendages are the means of locomotion of a cell through a fluid medium. Ameboid movement is suited only for movement on a solid. In contradistinction, the whiplike, or fishtail, beating of cilia and flagella effectively propel the cell in a fluid. But these structures have other functions as well. If the cell remains stationary, movement of cilia can serve to bring nutrients to the cell, and to remove waste products and foreign objects from the cell. The latter function, that of removing foreign objects, appears to be their sole purpose in multicellular organisms. For example, the nostrils are lined with ciliated

epithelium which serves to carry foreign objects, such as mucus droplets, away from the cell.

That cilia and flagella can provide a very effective means of locomotion is well illustrated by a consideration of the rate of movement. The speed varies with the type of cell and the environmental condition, but a rate of movement as high as 140 mm per minute has been recorded. This is to be compared with ameboid movement, which at a maximum, rarely exceeds 0.3 mm per minute.

Ciliary and flagellar movement are influenced by a variety of stimuli, but the mechanism essential to this type of movement is unknown.

SKELETAL MUSCLE

In the higher animal forms, the ability to change position is solely the function of skeletal muscle. But muscle is also essential for other vital processes. Cardiac muscle serves to pump the blood through the circulatory system. Smooth muscle is essential to bladder contraction, to the regulation of the size of the pupil of the eye, to activity of the gastrointestinal tract, as well as to many other visceral functions. The fundamental physiology of the three types of muscle in some aspects is identical, but in others there are important differences.

Chemistry of Contraction

When a muscle contracts, it performs work. The energy for this work evolves from a series of complex chemical reactions within the muscle. In some way this energy is utilized by the muscle fiber to bring about an alteration, or reorientation of protein molecules. It is this reorientation that causes a shortening of the muscle fiber.

It is believed that there are two distinct proteins involved in the shortening process: 1) myosin, and 2) actin. Myosin molecules are rod-shaped, approximately 1,500 Å long and 20 Å in diameter. The molecular weight is about 400,000. Actin is a much smaller molecule with a molecular weight of only 60,000. These proteins are thought to be arranged as shown in Fig. 6.3. Thus, instead of a folding of the molecules of the filaments to produce shortening as has been postulated by some, in this scheme there is a sliding together of the actin filaments over stationary myosin filaments. It is thought that there is some reaction between, or combination of, the actin and myosin

molecules to bring about this sliding movement. The nature of such bonding is not known. The conversion of adenosine triphosphate (ATP) to adenosine diphosphate (ADP) is the reaction that is believed to supply the energy for muscle shortening. But just how this reaction influences actin and myosin is not clear.

MYOSIN   ACTIN   2.5 μ
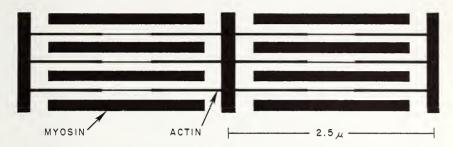
Fig. 6.3.   The Arrangement of Myosin and Actin in the Skeletal Muscle Fiber.

During maximal activity of a muscle the increase in rate of ATP utilization is far greater than the increase in rate of oxygen consumption by that muscle. Accordingly, unless there were a means other than oxidation, the supply of ATP would be quickly exhausted and the muscle would cease to contract. But ATP can be regenerated rapidly anaerobically by virtue of the interaction of ADP and phosphocreatine. Thus:

$$\text{Phosphocreatine} + \text{ADP} \longrightarrow \text{ATP} + \text{Creatine}$$

Just as ATP must be regenerated for continued muscle action, so must phosphocreatine be reformed from creatine and phosphate. This latter synthesis requires ATP. In short, both the energy of contraction and the energy for the reconversion of phosphocreatine stem from ATP. There must thus be another source of energy to synthesize ATP, other than the phosphocreatine reaction. This second source is supplied by the conversion of glycogen to lactic acid. In the intact organism, the lactic acid formed during contraction diffuses out of the muscle and is carried to the liver. In the liver, the lactic acid is converted to glycogen and then released into the blood as glucose. The glucose is carried back to the muscle and is once again converted into glycogen. There is thus the cycle shown in Fig. 6.4. This is sometimes referred to as the Cori cycle. The significance of

this cycle is that some of the energy for muscular contraction is supplied by metabolic processes which take place in the liver.

The conversion of lactic acid to glycogen requires oxygen. Lactic acid, itself, is oxidized and this oxidation provides the energy for the reconversion of the remaining lactic acid to glycogen. Actually, it is not a direct relationship. The oxidation of lactic acid to carbon dioxide produces energy for the reconversion of ADP to ATP. The ATP then supplies the energy for the conversion of the remaining lactic acid to glycogen. About one-sixth of the available lactic acid is used for the synthesis of ATP and five-sixths for the conversion to glycogen.
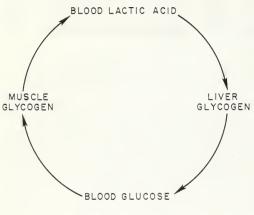
The present concept of the chemistry of muscular contraction, then, visualizes ATP as the central agent responsible for:

1. The actin-myosin interaction to cause shortening.

2. The synthesis of phosphocreatine from creatine and phosphate.

Fig. 6.4. The Cori Cycle.

3. The synthesis of glycogen from lactic acid.

The only reaction in which ATP does not participate is the conversion of lactic acid to carbon dioxide. This is an oxidative reaction, thus, ultimately contraction requires oxygen.

The interreactions deemed essential to muscular contraction are shown in Fig. 6.5. It should be noted that by the oxidation of lactic acid, energy is provided for the synthesis of ATP from ADP. ATP then provides the burst of energy needed for contraction, for the synthesis of phosphocreatine, and for the synthesis of glycogen.

Apparently all three types of muscle depend upon the ATP mechanism for contraction. There are, however, differences. In rapidly contracting skeletal muscle the immediate source of energy to maintain an adequate supply of ATP is the anaerobic transformation of glycogen to lactic acid. In slowly contracting smooth muscle, this

energy seems to be supplied mostly by the oxidation of fatty acids and acetoacetate. In cardiac muscle, ATP is maintained by the oxidation of acetoacetate and lactic acid. During exercise, skeletal muscle produces lactic acid which is carried by the blood to the liver, as described above. But some of this lactic acid also goes to the heart
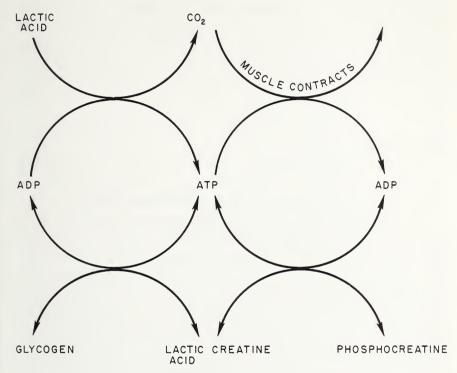
Fig. 6.5.   The Interreactions Essential to Contraction.

where it is oxidized to provide an additional source of energy for that organ.

Oxygen Debt

During vigorous exercise, in vertebrates at least, the oxidation of lactic acid and its conversion to glycogen do not keep pace with its formation. After the bout of exercise is completed, the organism continues to oxidize the accumulated lactic acid until the conversion to glycogen is complete. Oxygen, then, continues to be used for

muscle activity after the period of contraction is ended. In other words, during vigorous exercise muscles "live" beyond their means and, accordingly, incur an oxygen debt. This debt is paid back later, during the recovery period.

Isotonic and Isometric Contraction

When muscles are used for movement, they shorten. On the other hand, a muscle can contract without changing in length. In the first case the muscle shortens, but the tension of the muscle remains practically the same. Accordingly, it is termed isotonic contraction. In the second case, the length of the muscle stays practically the same, but the tension increases sharply. This is termed isometric contraction.

Heat Production

In isometric contraction all of the energy expended is recorded as heat. On the other hand, during isotonic contraction at least 25 percent of the energy expenditure appears as mechanical work.

The rate of heat production during and after contraction of a muscle varies. Two general phases of heat production have been described: 1) the **initial heat**, and 2) the **recovery heat.** Careful evaluation discloses that there is a burst of heat when a muscle contracts and another, but smaller, emanation, when it relaxes. Together they constitute the initial heat. But even after the muscle has relaxed completely it continues to produce heat due to the continuation of the chemical processes outlined above. This is the recovery heat. The recovery heat in the intact animal is less than in the isolated muscle because much of the lactic acid produced during contraction is removed by the circulation and carried to the liver for oxidation or transformation to glycogen.

Excitation

Skeletal muscle may be excited by direct stimulation; it rarely contracts spontaneously. In the intact organism all excitation is under the control of the nervous system. In man, hundreds, or even thousands of nerve fibers innervate each muscle. Usually, the nerve fiber arborizes into many branches each of which terminates on an individual muscle fiber. All of the muscle fibers thus controlled by a

single neuron constitute one **motor unit.** There is no protoplasmic continuity between the neuron and the muscle fiber. Rather a potential space remains between the two cell membranes. It is called the **myoneural junction.**

According to the most widely accepted current concept, the nerve impulse does not excite the muscle fiber directly. At the myoneural junction **acetylcholine** is liberated in response to the nerve impulse. Acetylcholine apparently alters the membrane of the muscle cell so as to permit an influx of sodium ions. This rapid influx creates a progressive movement of ions along the muscle fiber, that is, an impulse is propagated (see Chapter 7). In most mammalian muscles, the rate of propagation is about 3 meters per second.

Exactly how the muscle impulse initiates contraction is not known. In some manner the influx of sodium and the resulting imbalance triggers the ATP mechanism resulting in the interaction of myosin and actin.

Response to Single Stimulation

Figure 6.6 shows a typical response of skeletal muscle to a single stimulation. There is a measurable period of time between the stimulation and the onset of shortening. This delay is termed the **latent period.** The latent period in the gastrocnemius of the frog measures about 3.0 milliseconds. The latent period as measured in the student laboratory is usually longer than this because of the inertia in the measuring system. With better recording techniques it has been shown that the muscle, after stimulation, first relaxes somewhat, that is lengthens, before shortening. This is called **latency-relaxation.** The sequence, then, is a period during which no change occurs, followed by the latency-relaxation, followed by shortening. In most measurements, the latent period includes the latency-relaxation time.

Once the muscle begins to shorten it does so rapidly, quickly reaching a peak. This is the **contraction phase** which is followed by the **relaxation phase.** The duration of the entire response varies from one muscle to another. Leg muscle fibers exhibit a duration of from 50 to 100 milliseconds for a single twitch. In contrast, an eye muscle may complete its response in little more than 10 milliseconds.
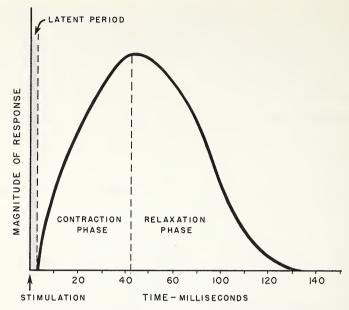
Fig. 6.6.   Response of Skeletal Muscle to a Single, Maximal Stimulus.

All-or-None Response

The magnitude of response of a single muscle fiber is independent of the type, or strength, of the stimulus. If the fiber responds at all, it responds maximally. Such a response is termed all-or-none.

Recruitment

A muscle is composed of many individual cells, or fibers. These cells have different thresholds for excitement. Accordingly, when a whole muscle is activated by a strong stimulus, the force of contraction is greater than when it is activated by a weak one. As the strength of stimulation increases, more and more muscle fibers respond. This progressive response of more units is termed recruitment. After all of the fibers are responding, an increase in stimulus strength can evoke no further increase in magnitude of response.

Initial Length

So long as conditions are maintained constant the response of a single muscle fiber is all-or-none. However, if conditions are altered,

the response may be different. For example, the force of contraction of a single fiber can be varied by changing the length of that fiber before it is stimulated, that is, the initial length. Figure 6.7 discloses that the magnitude of response increases proportionately with the initial length, up to a critical point, after which it decreases.
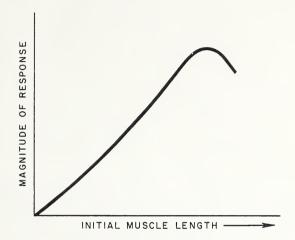
Fig. 6.7.   Influence of the Initial Muscle Length on the Magnitude of Response.

Refractory Periods

If a muscle is stimulated a second time and if that second shock quickly follows the first stimulus, there is no additional response. Since the muscle does not respond to a stimulus it is said to be refractory, and the period during which it will not respond is called the refractory period. In skeletal muscle it is extremely short being of the order of less than one millisecond. Actually, two refractory periods are recognized in this short interval: 1) the **absolute refractory period,** and 2) the **relative refractory period.** In the first, no response can be elicited no matter how strong the stimulus. But immediately following this brief interval it is possible to evoke a second response if a stimulus stronger than the threshold stimulus is used.

Response to Repetitive Stimulation

Figure 6.8 shows that it is possible to cause skeletal muscle to contract a second time before it has relaxed from the first contraction.

It will also be noted that the force of contraction achieved by the second contraction is greater than it is for a single twitch. In effect, the second response has been added to the first. This addition of contraction waves is known as **summation.** The increased force of contraction of a whole muscle with increased strength of stimuli is due to recruitment, that is, more fibers respond. But summation occurs in a single muscle fiber. Yet this is not a violation of the all-or-none law, for the conditions under which the fiber contracts the first and second time are different. In the first contraction, the fiber is completely relaxed; in the second, it is partially contracted. Actually, the amplitude of the second contraction is less than that of the first, but it has been added to that of the first, thus the total amplitude, that is the sum of the two, is greater than either one. In short, recruitment is the addition of motor fibers; summation is the addition of contraction waves of individual fibers.

It is possible to stimulate a skeletal muscle fiber so rapidly that there is no time between stimuli for the fiber to relax. The response, as seen in the lower portion of Fig. 6.8, is smooth, sustained, and maximal. Such a response is termed **tetanus.** It will be noted that although the response is smooth and appears to be but one prolonged single contraction, the recording of the muscle action potentials reveals that there is an individual response for each stimulus.

Finally, Fig. 6.8 shows that a sustained response may be obtained with somewhat less rapid stimulation but in this case the contraction is not smooth. Individual waves may be discerned. This type of contraction is termed **incomplete tetanus.**

If a muscle is stimulated repetitively for a long enough period of time, one notes that the force of contraction progressively decreases and finally a point is reached at which no contraction occurs. The muscle is said to be **fatigued.**

Tonus

A maximal response of a muscle is obtained when all of the component fibers act synchronously. A weaker but still sustained contraction is possible when fewer fibers contract asynchronously. Because any one fiber only contracts for a brief period of time before its work is taken over by another fiber, fatigue does not occur. Except
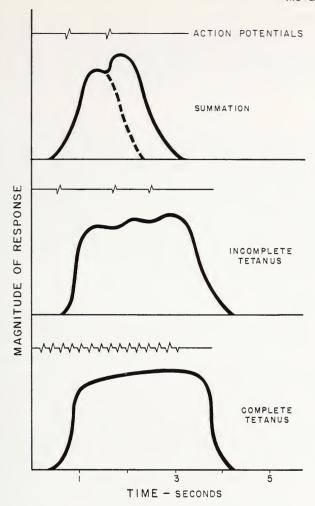
Fig. 6.8.   Response of Skeletal Muscle to Repetitive Stimuli.
Repetitive stimulation causes summation, incomplete tetanus,
or tetanus, depending upon the frequency of stimulation.

when a muscle is completely relaxed, as may occur during sleep, most
skeletal muscles have at least a part of their fibers in the state of
partial contraction. The muscle thus feels firm, it resists being
stretched, it exhibits tension. The muscle is said to possess tone,
which means tension.

Tonus of skeletal muscle depends completely upon its innervation.

If it is denervated it has no tone. It is **atonic.** Tonus is best defined as involuntary resistance to passive stretch. This resistance is due to a reflex response initiated by stretch. If the reflex is extremely sensitive the response will be rapid and vigorous. The muscle is **hypertonic.**

Paralysis

Skeletal muscles can be contracted voluntarily. This response to the will is mediated by the central nervous system and the motor nerves that innervate the muscles. If there is damage to the central nervous system, or to the motor nerve so that no impulses reach the muscle, that muscle can no longer respond to the will and it is said to be paralyzed. This does not mean that the muscle is incapable of responding. A direct stimulation of the muscle, say by an electrical current, will still evoke contraction.

Growth and Waste

The growth of a skeletal muscle cell is controlled by many intracellular processes just as the growth of any cell is (see Chapter 8). But in addition, the size of a skeletal muscle cell may also be influenced by the use of that fiber. Thus, if a muscle is used to perform heavy work it will enlarge beyond its normal size. The enlargement of a cell in response to increased activity is termed **hypertrophy.** Conversely, a muscle that is not used decreases in size. This is termed **atrophy.** In some cases a whole muscle increases in size not by the growth of the individual fibers, but due to the increased number of fibers. Such growth is referred to as **hyperplasia.**

CARDIAC MUSCLE

It has been pointed out that the chemistry of contraction of cardiac muscle is almost identical with that of skeletal muscle insofar as it is known. The only difference seems to be that in cardiac muscle, ATP is maintained by the oxidation of acetoacetate and lactic acid; whereas, in skeletal muscle this energy comes from anaerobic transformation of glycogen to lactic acid. A portion of this lactic acid is oxidized by the heart.

All-or-None Response

Because cardiac muscle is a **syncytium**, that is, there is protoplasmic continuity between the fibers, the entire muscle responds in an all-or-none manner. Accordingly, a graded response due to recruitment is impossible.

Starling's Law of the Heart

The force of contraction of cardiac muscle, just like skeletal muscle, varies with the initial length of the fiber. In the heart the length of the fiber immediately before contraction is determined by the volume of blood in the ventricle. On the other hand, the force of contraction of the ventricle is manifested by the volume of blood ejected during the contraction phase. Thus, according to Starling's law, the volume of blood ejected by the ventricle is proportional to the volume of blood in the ventricle before contraction. In short, the more blood that enters the heart, the more that is pumped out. This is true, however, only up to a definite point, after which the amount of blood ejected decreases with increased volume of inflow.

Refractory Periods

The refractory periods of cardiac muscle, both absolute and relative, are considerably longer than they are for skeletal muscle. In cardiac muscle the absolute refractory period extends throughout the contraction phase, and the relative refractory period persists almost until the muscle is completely relaxed. In skeletal muscle it is definitely advantageous to have a smooth, maximal, sustained contraction, that is, tetanus. In the heart, such a contraction would be fatal. In order to pump, the muscle must contract to eject blood and then relax to permit another volume to enter the ventricle. Because of the long refractory periods cardiac muscle cannot be thrown into tetanus.

Inherent Rhythmicity

If skeletal muscle is denervated, it is paralyzed. It cannot be contracted voluntarily, and it does not contract spontaneously. Cardiac muscle is not under the control of the will, and although it is regulated by its innervation, it will continue to contract rhythmically

after denervation. Cardiac muscle has an inherent rhythmicity. The rhythm may be altered by changing the temperature or other environmental factors. The mechanism responsible for this inherent rhythmicity is not understood.

Regeneration

Cardiac muscle hypertrophies in response to an abnormal work load. But unlike skeletal and smooth muscle, cardiac muscle does not regenerate if it is injured. Instead of new muscle, the injured area becomes a scar consisting of connective tissue. Thus, following a coronary occlusion, an area of cardiac muscle may die because of oxygen deprivation. This area then becomes a scar, called an **infarct,** and no longer can contribute to the force of contraction.

SMOOTH MUSCLE

The physiology of smooth muscle is very inadequately understood. This is due not to a lack of interest in the problem, but rather to the fact that smooth muscle rarely responds in a consistent manner even when conditions are rigidly controlled.

Contraction

The contraction of smooth muscle differs markedly from that of skeletal muscle. The response is considerably slower, and following a single stimulus, the contraction may persist for several seconds. Smooth mucle is capable of spontaneous contraction (Fig. 6.9). Even when it is completely denervated, it often continues to shorten and relax with variable rhythm.

The energy of contraction is supplied by reactions similar to those described for skeletal muscle. However, for the slow contractions of smooth muscle the oxidation of fatty acids and acetoacetate seems to suffice to provide the energy to maintain an adequate supply of ATP.

Accommodation

The more tension placed upon skeletal and cardiac muscle, the greater will be the resistance to being stretched. This is not true of smooth muscle. After smooth muscle is stretched, it seems to relax

and lengthen so that the resistance remains about the same. In the urinary bladder which is made up of smooth muscle, as urine accumulates the muscle is stretched. But the pressure within the bladder remains practically unchanged. This ability of smooth muscle to relax in response to greater stretch is termed accommodation. In other words, a larger volume is accommodated with little change in pressure or tension.
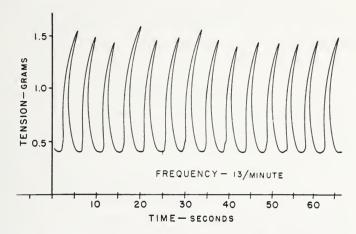
Fig. 6.9. Spontaneous Contractions of Smooth Muscle (Rabbit Gut).

SUMMARY

Four types of movement are recognized: 1) **protoplasmic streaming,** 2) **ameboid,** 3) **ciliary** or **flagellar,** and 4) **muscular.** In protoplasmic streaming the cell does not move, only the protoplasm. In ameboid movement the protoplasm moves but so does the cell. Ciliary and flagellar movements serve to propel the cell through a fluid medium. These primitive types of movement are extremely susceptible to environmental conditions.

The shortening process, characteristic of muscle, is believed to depend upon the interaction of two proteins, **myosin** and **actin.** The energy for this process comes from the phosphate bonds of **ATP.** ATP also provides the energy for the synthesis of **phosphocreatine** from **creatine** and **phosphate,** and the synthesis of **glycogen** from **lactic acid.** In the absence of oxygen the breakdown of phospho-

creatine provides the energy for resynthesis of ATP. Another source of energy for ATP resynthesis is the degradation of glycogen to lactic acid. When oxygen is available, lactic acid is oxidized to provide the energy for the other reactions.

Skeletal muscle, because of the anaerobic reactions, can accumulate an **oxygen debt.** In **isotonic** contraction the muscle shortens but the tension remains unchanged. In **isometric** contraction the muscle does not shorten, but the tension increases. When muscle contracts there is a burst of heat, termed the **initial heat.** There is further heat generated during recovery, the **recovery heat.**

Skeletal muscle contracts in response to direct stimulation, or by activation of its innervation. At the **myoneural junction** acetylcholine is liberated. This compound activates the muscle. Following stimulation there is a **latent period** which includes the **latency-relaxation.** This is followed by the contraction and relaxation phases. A single muscle fiber has an **all-or-none** response. The addition of more muscle fibers to the total response is termed **recruitment.** The force of contraction varies with the **initial length** of the fiber. Immediately following one stimulation the muscle will not respond to a second. This period of time is termed the **refractory period.** If skeletal muscle is stimulated repetitively the contractions summate, and if the rate of stimulation is great enough, **tetanus** results. The resistance to passive stretch is termed **tonus.** Denervated skeletal muscle is **paralyzed;** it will then **atrophy.**

Cardiac muscle is a syncytium and therefore the muscle as a whole exhibits an all-or-none response. **Starling's law of the heart** states that the volume of blood ejected is proportional to the volume of blood in the ventricle before contraction. This response of more vigorous myocardial contraction is a manifestation of the influence of initial length. The refractory periods of cardiac muscle are long and thus this type of muscle will not go into tetanus. Cardiac muscle has an **inherent rhythmicity.**

Smooth muscle contracts slowly. Oxidation of fatty acids and acetoacetate seems to suffice to provide energy to resynthesize ATP. Smooth muscle has an inherent rhythmicity and does not depend upon its innervation. In response to stretch, this type of muscle undergoes **accommodation.**

Problems

1. Outline the interreactions essential to the chemistry of contraction.
2. What is the limiting factor to the accumulation of an oxygen debt?
3. Why does the recovery heat differ in a muscle in the intact organism from that recorded in an isolated muscle?
4. List the major differences between cardiac, smooth, and skeletal muscle.
5. Define:

a. Recruitment	d. Summation	g. Paralysis
b. Latent period	e. Tetanus	h. Hyperplasia
c. Refractory period	f. Tonus	i. Accommodation

Additional Reading

Drummond, G. I. and Black, E. C., "Comparative Physiology: Fuel of Muscle Metabolism," *Ann. Rev. Physiol.,* **22,** 169-190 (1960).

Dubuisson, M., "Muscular Contraction," Charles C Thomas, Springfield, Illinois, 1954.

Karpovich, P. V., "Physiology of Muscular Activity," 5th ed., W. B. Saunders Co., Philadelphia, Pa., 1959.

Szent-Gyorgyi, A., "Chemical Physiology of Contraction in Body and Heart Muscle," Academic Press, Inc., New York, N. Y., 1953.

Weber, H. H., "The Motility of Muscle and Cells," Harvard University Press, Cambridge, Mass., 1958.

ELECTRICAL ACTIVITY

ONE OF THE characteristics of the cell membrane, discussed in Chapter 3, is its ability to maintain an unequal distribution of ions between the inside and the outside of the cell. Due to this ionic imbalance a potential difference exists between the inside and outside of the cell. Because the potential derives from the membrane characteristic, it is termed the **membrane potential.** This is a property of all living cells. When nerve and muscle cells become active, that is, when nerve propagates an impulse and muscle contracts, there is associated with these functions electrical activity. The electrical activity is manifested by a potential difference on the surface of the cell between the active and nonactive area. Because these electrical potentials are associated with cell activity they are termed **action potentials.**

A word on terminology is in order before beginning the discussion of potentials. It has become the practice to speak of impulse "conduction" when referring to nerve and muscle. This is not, strictly speaking, correct. A metal wire conducts an electrical current by the transfer of electrons. Likewise it is possible to conduct an impulse along the wet surface of living tissue. But the mechanism to be discussed in this chapter is primarily concerned, not with conduction in the sense of electron transfer, but rather with **propagation,** a term that conveys the self-generating process characteristic of action potentials. In brief, inanimate substances conduct an impulse; living cells propagate an impulse.

MEMBRANE POTENTIALS

If one electrode is placed on the surface of a cell and the other within a cell, a current flows (Fig. 7.1). As has already been mentioned, this is a manifestation of the membrane potential. But because the potential exists when the cell is at rest, the membrane potential is also known as the **resting potential.**
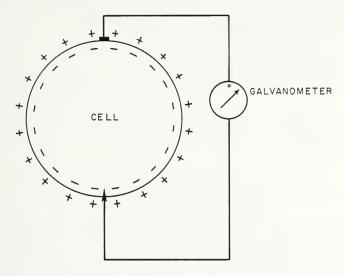
Fig. 7.1. Membrane Potential. When one electrode is within a cell and the other on the surface, a current flows because of the potential difference.

Origin of the Potential

Quite obviously, if a membrane existed which restricted cations to a solution on one side and anions to a solution on the other side, there would be a potential difference between the two solutions. And if electrodes were placed in the solutions and connected with a wire, a current would flow because of the potential difference (Fig. 7.2). This situation may be approached by the use of a membrane that is permeable to one ion but not the other. For example, assume that the membrane permits potassium ions to pass, but not chloride ions. If a higher concentration of KCl is placed on one side than on the other, there will be diffusion of K ions from the high concentration to the low. But the Cl ions cannot go along. Therefore, at

equilibrium there will be more positive ions on the low concentration side and more negative ions on the other. A potential difference between the two solutions will now exist.

Until fairly recently it was thought that the cell membrane was relatively impermeable to certain ions. Radioisotope studies have shown conclusively that at least insofar as sodium, potassium and chloride ions are concerned, they can diffuse through living membranes. Yet, by some poorly understood mechanism, the ionic imbalance shown in Table 7.1 is maintained.

TABLE 7.1. Intra- and Extracellular Ions *

Ion	Intracellular mEq/l	Extracellular mEq/l
Potassium	120	5
Sodium	5	145
Chloride	5	105

* Values in this table are merely representative. They vary considerably from cell to cell and species to species.

It should be emphasized that the values given in Table 7.1 are simply average, representative figures. It is extremely difficult to determine intracellular concentrations because the methods available probably alter the membrane sufficiently to disrupt the normal equilibrium state. In addition, the values seem to vary with different types of cells. The concentration of potassium within the cell, for example, may be 20 to 50 times greater than it is in the external environment, depending upon the type of cell analyzed.

The relative rate of diffusion of an ion through a membrane is often spoken of as **conductance**. A better term is **mobility**. The mobility of potassium by the resting membrane is far greater than it is

Fig. 7.2. Separation of Ions Causing Current Flow.

for sodium. In addition, sodium that does enter the cell is quickly expelled by what has been termed a sodium pump. Because of these two factors, it has been postulated that it is the difference between the potassium concentration on the inside and outside of the cell that determines the magnitude of the resting potential. Due to the higher concentration of potassium on the inside there is a tendency for the ion to diffuse out. This, however, creates internal negativity which opposes the outward diffusion of the ion. Ultimately, an equilibrium is reached. Clearly, if the potassium concentration of the interior of the cell is increased, a greater resting potential should result. Conversely, if potassium is added to the surrounding fluid the potential should decrease. These theoretical conclusions have been confirmed experimentally.

If the imbalance of potassium is the true cause of the membrane potential, the potential may be calculated using the Nernst equation:

$$E = \frac{RT}{nF} \log_{10} \frac{C_1}{C_2}$$

where: E = the potential difference in millivolts
T = the absolute temperature
R = the gas constant
F = the Farady
n = the valence change
C_1 = the concentration inside the cell
C_2 = the concentration outside the cell

Substituting appropriate values the equation becomes:

$$E = 61 \times \log \frac{C_1}{C_2}$$

when the temperature is that of mammals, namely 37°C. If the values for potassium given in Table 7.1 are now substituted in the equation, the resting potential may be calculated:

$$E = 61 \times \log \frac{120}{5}$$
$$= 61 \times \log 24$$
$$E = 84 \text{ millivolts}$$

Actual measurements of the resting potential agree remarkably with such calculated values. But even though there is agreement between the actual potential and the calculated values, this does not prove that the Nernst equation, taking into consideration the moment of but one ion, is valid. The better concept would seem to be that the imbalance of all the major ions present in the intracellular fluid contributes to the membrane potential.

Polarization and Depolarization

When a resting potential exists, the membrane is said to be polarized, or charged. In other words, there is an ionic imbalance between the solutions on either side, and, as explained above, there are more cations on one side and more anions on the other. The cell membrane, then, is visualized as being polarized with negative ions on the inside and positive ions on the outside.

If the membrane becomes more permeable so that there is a free diffusion of all ions, a balance will soon be reached in which the concentration of ions on both sides will be the same, and, in each solution, the number of anions and cations will be the same. Accordingly, there will no longer be a potential difference between the two solutions, and the membrane will no longer be charged, it will be depolarized.

There are many ways in which a cell may be depolarized. In all cases, the underlying mechanism is an increase in membrane permeability. Thus, mechanical stimuli, such as crushing, disrupts the integrity of the membrane, permits a flow of all ions in both directions and therefore the membrane becomes depolarized. Various chemical agents have the same effect. Of these, acetylcholine is perhaps the most important because this is a compound that can be produced by nerve cells. It thus serves to fire, that is, to depolarize other nerve cells, or muscle fibers. As will be discussed presently, acetylcholine increases the membrane permeability thereby permitting the influx of sodium ions and the efflux of potassium ions. And finally, electrical stimuli are also capable of depolarizing the cell. In this connection it should be noted that if a positive electrode is placed on the outside of the cell and the negative electrode inside, the membrane will not be depolarized. On the contrary, it becomes

hyperpolarized. But if the electrodes are reversed, then the membrane is depolarized.

Reversal of Potential

It is conventional to consider the resting potential as being negative, thus, in the calculation above, the cell would have a potential of −84 millivolts. When the membrane is depolarized, it would be expected that the potential would simply return to zero. But as can be seen in Fig. 7.3, there is a reversal, that is, not only does the poten-
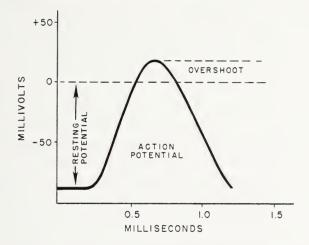
Fig. 7.3. Alteration in Resting Potential Resulting from Membrane Depolarization.

tial difference become zero, it sometimes goes beyond and becomes positive. This is termed the **overshoot**. The normal resting cell membrane is visualized as being negative on the inside and positive on the outside. Apparently, during depolarization the inside may become positive for a brief interval.

It will be recalled, that although there is a great difference between internal and external sodium ion concentration, the mobility for sodium is very low, being on the order of some 50 times less than the mobility for potassium. Thus, it has been concluded that it is primarily the outward diffusion of potassium ions that creates negativity within the cell. Ultimately the internal negativity becomes great enough to prevent further outward diffusion of potassium ions so

that an equilibrium is reached and the membrane is polarized. However, an adequate stimuli changes the permeability characteristics of the membrane. This greatly increases the mobility for sodium. Sodium ions now diffuse into the cell rapidly. Because of this influx the inside ultimately becomes positive to the outside, and the potential is reversed from negative to positive.

An important component of this theory of depolarization is that as sodium ions move into the cell the membrane becomes more permeable so that the sodium ion mobility is increased. Accordingly, once initiated the mechanism is self-regenerative. In short, an adequate stimulus depolarizes the membrane sufficiently to permit sodium ions to diffuse inward. The inward diffusion of sodium ions further increases membrane permeability so that the rate of sodium ion diffusion becomes even greater. Consequently, there is a great surge of positive ions into the cell. Figure 7.3 shows that the upward limb of the recorded potential begins slowly, then rapidly rises almost vertically. This is a manifestation of the self-regenerative process.

It has been calculated that the complete inward diffusion of sodium ions would produce a potential of about 115 millivolts. In other words, the resting potential should go from —84 to zero and then on to +31, a total of 115 millivolts. And this would occur if sodium ions alone moved. The reason the potential, upon depolarization, does not reach this magnitude is because the changed membrane permeability also increases potassium mobility, but this increase is delayed. When it does occur, the outward diffusion of potassium ions checks the potential created by the inward diffusion of sodium ions so that the potential now decreases and the mobility of sodium ions progressively diminishes. Sodium is once again pumped out so that the resting state is reattained.

The complete sequence of events, then, is thought to be as follows: In the resting, or polarized state, the mobility of sodium is so low, in comparison to the mobility of potassium, that it is the outward diffusion of potassium that creates the resting potential. The stimulus changes membrane permeability so that the mobility of sodium increases. This ion rushes into the cell and reverses the polarity. The mobility of potassium now increases and potassium diffuses out to offset, in part, the sodium effect. By some mechanism, not clearly

understood, the sodium is pumped out, the potassium relationship reestablished, and the membrane is once again polarized.

It must be emphasized that the values presented here are average, representative figures. Different cells with different ion concentrations at various temperatures will, of course, have a different set of values for the potentials. The values of potentials usually given are those obtained using molluscan (squid) nerve. Reliable values for mammalian nerve are difficult to obtain.

Metabolic Demands

If the cell is deprived of oxygen, the membrane potential decreases, and in many cells, completely disappears. This indicates, that in the living cell, the permeability characteristics essential to the maintenance of the potential depend upon metabolic processes. The specific reactions that are essential to membrane integrity are not known. Apparently glycolysis plays a role. This is shown by the fact that the membrane potential is decreased by chemical agents that inhibit glycolytic reactions. It is probable that ATP provides the energy required for membrane permeability, and perhaps for the sodium pump, if such a pump in fact exists. As has already been outlined, ATP synthesis requires energy which may be supplied by various reactions, including glycolysis. Ultimately, however, oxygen must be made available.

ACTION POTENTIALS

As has just been explained, there is a potential difference between the inside and outside of the cell. When the cell becomes active, there is a marked change in that potential; this change, then, may be thought of as the action potential. There are various ways to record the action potential. The easiest method is to place the two recording electrodes on the surface of the cell, usually a nerve or muscle cell (Fig. 7.4). At rest there is no potential difference between the two electrodes. However, when the area of the cell under one of the electrodes becomes active, that area is depolarized, and, because of the influx of sodium ions, becomes negative in relation to the surface of the cell under the other electrode. There is now a potential difference.

Diphasic and Monophasic Action Potentials

As shown in Fig. 7.4, if the two recording electrodes are placed upon the surface of a nerve cell which is then stimulated to propagate an impulse, the resulting recording will be of a diphasic action potential. That is to say, there is a deflection in one direction when the area under one electrode is depolarized, and then there is a deflection in the opposite direction when the area under the second electrode becomes depolarized. By the time depolarization under the second electrode occurs, the area under the first is repolarized, resulting in a deflection in the opposite direction. However, when the impulse is between the electrodes, both areas have the same polarity and thus there is no deflection.
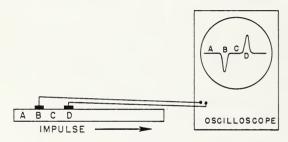
Fig. 7.4. Action Potentials as Visualized Using an Oscilloscope. A diphasic action potential is seen because of the reversal of charge as the membrane becomes depolarized first under one electrode and then under the other.

It is often desirable to eliminate the second deflection, that is, to obtain a recording with but one deflection, a monophasic action potential. This can be achieved in at least three ways: 1) with one electrode inside and the other on the outside of the membrane, a single deflection from the resting position will result when the membrane is depolarized; 2) with the two electrodes on the surface of the nerve and an area between the two electrodes blocked by an anesthetic or by injury; and 3) with one of the electrodes on an injured area (Fig. 7.5).

Components of an Action Potential

Assume that the second method of recording a monophasic action potential is used, that is, with an area of the nerve between the two
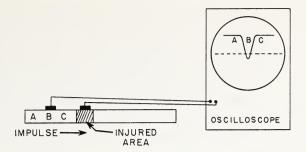
Fig. 7.5. Monophasic Action Potential. At rest the injured area is negative in relation to the uninjured area of membrane. Depolarization (B) brings the potential toward the isoelectric line, and it may become negative in relation to the injured area (overshoot).

electrodes blocked. At rest (Fig. 7.6), there will be no potential difference between the two surface electrodes. Now the nerve is stimulated so that the area under the first electrode is depolarized. This area becomes negative, thus Fig. 7.6 shows a downward deflection as the area under the first electrode becomes depolarized. The slope of this deflection is very steep and the change, after the peak is reached is so fast that a point, or spike, is described. Accordingly,
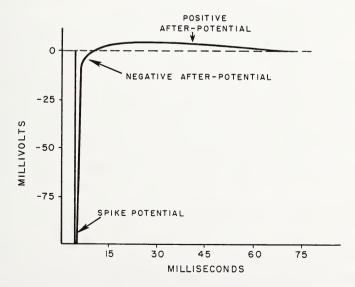
Fig. 7.6. Components of the Action Potential.

this part of the action potential is termed the **spike potential.** The magnitude of the spike potential, as already discussed, is usually somewhat greater than the resting potential. The actual magnitude varies with the type of cell, the temperature, and other factors.

In some cases the descending limb of spike breaks rather abruptly so as to widen the base. The part of the action potential is termed the **negative after potential.** Following this, the curve may rise above the isoelectric line, that is, to become positive. This is the **positive after potential.**

Excitability Thresholds

During the period of time required for the upswing of the spike potential, the cell will not respond to another stimulus no matter how intense. This is the **absolute refractory period.** In mammalian nerve cells it has a duration of about 0.4 milliseconds. It is followed by the **relative refractory period,** during which an above-threshold stimulus will evoke a response. There is then a sudden reversal of excitability. During the period of the negative after potential the excitability threshold is lower than normal so that, at this time, the nerve cell may be more easily activated. Another reversal occurs during the positive after potential period in which the excitability threshold is once again higher than normal.

Ionic movement has been used to explain the resting and action potentials. Although there is still not general acceptance of this hypothesis, nonetheless it does accord remarkably well with experimental data. And it does explain better than any other hypothesis the after potentials and the change in excitatory thresholds.

The absolute refractory period limits the number of impulses that can be propagated per unit time, that is, the frequency of response. If the nerve cell has an absolute refractory period of 0.4 milliseconds then the maximum impulse frequency cannot exceed 2.500 per second.

Propagation of the Action Potential

In order to initiate the action potential there must be sufficient alteration in the membrane so as to permit the inward diffusion of sodium. When this occurs, a potential difference develops between the depolarized area and the contiguous regions. Accordingly, current

flows. This current flow, on either side of the depolarized area, is adequate to depolarize the contiguous membrane sufficiently to permit the initial influx of sodium. This is a self-regenerative process so that from the original point of stimulation waves of depolarization sweep along the cell in both directions.

If a nerve cell is stimulated at one end, the wave of depolarization will then be propagated from that end to the other. Since this process is associated with a progressive depolarization of the membrane, there is a wave of negativity that moves along the nerve cell. In other words, an **impulse** is propagated.

Velocity of Propagation

The velocity at which a nerve cell propagates an impulse is a function of its diameter. The fibers of greatest diameter have been found to propagate impulses at a velocity of up to 150 meters per second. Very small mammalian nerve cells may have a propagation velocity of less than 1 meter per second. By way of contrast, mammalian cardiac muscle propagates impulses at approximately 0.5 meters per second.

All-or-None Principle

Every cell has a specific threshold of excitation. But once that threshold is exceeded the response of the cell is unchanged by increasing the intensity of the stimulus. This all-or-none principle was mentioned in relation to the contraction of muscle. It also applies to the propagation of an impulse. A threshold stimulus will evoke an action potential of a specific magnitude, configuration, and duration. The action potential does not change in any of these values when the stimulus intensity is increased. The action potential is then propagated the length of the cell at a specific velocity. This velocity is a property of the cell and is unchanged by altering the stimulus. The number of impulses propagated by a cell per unit time, however, is directly proportional to the rate of stimulation up to a maximum beyond which the cell will not respond.

Chronaxie

The chronaxie is really a measure of the irritability of the cell. Irritability means the ability to respond to a change in the environ-

ment. Typical responses are the propagation of an impulse in a nerve, contraction of a muscle, or secretion of a gland. The chronaxie can thus be used to compare the responsiveness of various cells. Any manifestation of response may be used, but generally it proves most satisfactory to use the action potential as an indication of response.

In order to calculate the chronaxie, a **strength-duration curve** is constructed. As the term indicates, there are two important aspects of an electrical stimulus: 1) the strength usually measured in millivolts, and 2) the duration generally measured in milliseconds. If a nerve is to be evaluated in this way, stimulating and recording electrodes are placed on the cell. Then a stimulus duration of, for example, 2 milliseconds is selected and the voltage increased until a response, as manifested by an action potential, is obtained. The duration is then decreased and the procedure repeated. In this way, the curve depicted in Fig. 7.7 is obtained.

It can be seen that the weaker the stimulus, the longer it must act on the cell in order to evoke a response. The **rheobase** stimulus has an intensity just adequate to evoke a response when applied for an infinite time. In practice, the part of the curve parallel to the hori-
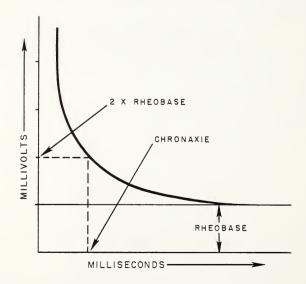
Fig. 7.7. Strength-Duration Curve from Which Chronaxie May be Calculated as Indicated.

zontal axis is extrapolated back to the vertical axis and the voltage read. This is the rheobase. The rheobase voltage is doubled and from the curve the time required for a response using a stimulus of this intensity is ascertained. This is the chronaxie. It is expressed in units of time, usually milliseconds. By definition, then, the **chronaxie** is the time a stimulus of an intensity twice the rheobase must be applied to a cell to evoke a response.

Action Potentials in Other Cells

Potential differences have been found to be associated with a wide variety of cells. The discussion in this chapter was concerned, for the most part, with electrical activity in nerve cells because knowledge of this type of cell is far greater than for any others. It has already been mentioned that there is an action potential associated with the contraction of muscle. Figure 7.8 shows that after a muscle is stimulated, the action potential takes place before the muscle actually shortens. This is true of all three types of muscle. The action potentials associated with heart contractions are widely used in clinical medicine. They form a pattern that is termed the **electrocardio-**
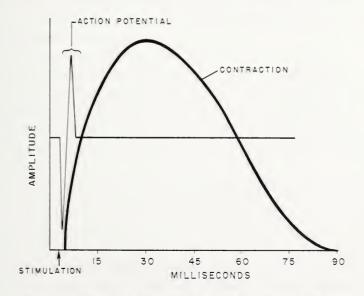
Fig. 7.8. Temporal Relationship of the Muscle Action Potential and Contraction.

gram (Fig. 7.9). If the electrodes are placed relatively close together on one part of the heart, a typical diphasic action potential will be recorded. The record seen in Fig. 7.9, however, was obtained with one electrode on the right arm and the other on the left leg. Impulses from the heart are conducted by the tissues of the body to these limbs. The complex pattern seen in the normal electrocardiogram then represents the resultant of all of the electrical activity that occurs in the heart during its contraction and relaxation phase. Accordingly the record is composed of several deflections during each beat of the heart (see Chapter 12).

Glandular cells, as is the case with all other cells, exhibit a membrane potential. When the gland is stimulated, it secretes, and associated with this activity there is a change in the membrane potential. Thus it can be said that as with nerve and muscle, glandular activity is also accompanied by an action potential.

Fig. 7.9. Electrocardiogram. Action potentials from the beating heart can be recorded by placing electrodes on the extremities.

The giant nerve fiber of the squid has been extensively used for the study of electrical activity. The advantage of this organism is that an electrode may be easily placed within the nerve cell. At rest these cells have a potential difference between the inside and outside of the membrane of about 50 millivolts with the inside negative to the outside. During activity the polarity changes so that the outer surface now is more than 50 millivolts negative to the inside, a total change of at least 100 millivolts.

The ubiquitousness of electrical activity in living cells is demonstrated not only by the presence of resting and action potentials throughout the animal kingdom, but in plant cells as well. For example, in Nitella, an alga, stimulation causes the spread of an action potential at a rate of about 1 centimeter per second. This is a propagation velocity that compares with small mammalian nerves. Action potentials have been recorded from more complex plant forms. Light causes a reaction in some cells and when this occurs the lighted side of the leaf becomes positive to the other side. Simply pulling on a

plant often suffices to create an action potential. The interesting aspect is that the stimulated area becomes negative. As has been emphasized, it is generally true that the area of stimulation of most all cells becomes negative in reference to the unstimulated parts. Action potentials have also been recorded from the surface of the fascinating insectivorous Venus fly-trap. When this plant is stimulated to move, there is a typical action potential.

Electric Organs

The electrical activity of living cells at rest, or during activity, is generally of the magnitude of about 100 millivolts, that is, a tenth of a volt. Some organisms, however, possess electric organs which are capable of generating voltages of over 600 volts with a power output of about 100 watts. This electrical activity is used for defensive purposes and it also serves to capture prey. Various fishes and elasmobranchs, have electric organs. The electric eel, found mostly in the Amazon, has been extensively studied.

The electric eel, *Electrophorus electricus,* possesses four electric organs. They occupy the major part of the elongated body of this species. The electric organs are divided into compartments that have been termed **electroplaxes.** Apparently the electroplaxes are activated by their innervation.

The electroplax is really modified skeletal muscle which has retained its innervation. It responds in much the same way that muscle does; it fatigues, and transmission of the impulse from the nerve to the electroplax may be blocked by the same chemicals that block transmission from nerve to muscle. Even more convincing are the intra- and extracellular potassium concentrations and the magnitude of the resting potential. These are very similar to muscle. Thus, it has been found that the electroplax has a resting potential of about 85 millivolts. Upon stimulation there is a reversal of polarity so that the inside now becomes positive to the outside with a potential difference of close to 65 millivolts. This means that there is an action potential of approximately 150 millivolts. This is the same as for skeletal muscle. One then wonders why an electric organ can generate high voltages whereas a muscle cannot. The explanation seems to lie in the "wiring." In a muscle the individual cells are connected, or wired, in parallel. But in the electric organ, the individual elec-

troplaxes are connected in series so that the voltages summate. The maximum voltage produced by an electric organ depends upon the number of electroplaxes. In the electrical eel, each organ has about 5,000. If each electroplax can generate 150 millivolts, the organ could, theoretically, produce a potential difference of 750 volts. This agrees well with observed voltages. That the number of electroplaxes is the determining factor is also shown by the observation that the electric ray has only about 60 electroplaxes in its electric organ and it can generate a potential difference of less than 10 volts. Such low voltages may serve as a direction finding device, rather than for protection or capture of prey.

SUMMARY

The potential difference between the inside and outside of a resting cell is termed the **resting,** or **membrane potential.** The rate of diffusion of an ion through a membrane is a measure of its **conductance** or **mobility.** Because the mobility of potassium through the resting cell membrane is far greater than is the mobility of sodium, and because sodium is quickly expelled from the cell, it is thought that the difference between the potassium concentration on the inside and outside of the resting cell primarily determines the magnitude of the resting potential. The cell membrane at rest is said to be **polarized.** Upon appropriate stimulation there is an inrush of sodium ions so that the membrane becomes **depolarized.** In most cells, the polarity is briefly reversed. Polarization and the maintenance of the resting potential depends upon intracellular metabolic processes. Deprivation of oxygen, or injury to the cell, leads to depolarization.

The change in membrane potential that occurs upon stimulation of the cell is termed the **action potential.** The membrane becomes progressively depolarized in both directions from the point of stimulation so that an **impulse** is propagated. The velocity of propagation varies from cell to cell. In large mammalian nerves it may be as great as 150 meters per second. The rapid changes in potential form the **spike potential.** In some cells this is followed by the **negative** and **positive after potentials.** The after potentials, the change in excit-

ability thresholds and the all-or-none reaction are explicable by the ionic hypothesis.

Responsiveness of a cell is indicated by its **chronaxie** which is defined as the time a stimulus of an intensity twice the rheobase must be applied to a cell to evoke a response.

It is probable that all cells exhibit an action potential upon activation. They have been recorded in nerve, muscle, and glands, as well as in plants. Some fishes and elasmobranchs possess **electric organs** which are capable of generating voltages of over 600 volts. This is accomplished by many individual **electroplaxes** being connected in series. The voltage generated may be used for protection, capture of prey, or as a direction-finding device.

Problems

1. Explain the ionic hypothesis of membrane and action potentials.
2. How does ionic movement explain the after potentials, the change in excitability thresholds, and the all-or-none response?
3. Assume that the inside of an electroplax has a potassium concentration of 140 mEq/liter and the extracellular fluid has a potassium concentration of 4 mEq/liter. When the electroplax is discharged there is an overshoot of 45 millivolts. If there are 3,000 electroplaxes in the electric organ, what would be the maximal voltage that could be generated?
4. Define:

    a. Chronaxie   c. Conductance
    b. Rheobase   d. Depolarization

Additional Reading

Brazier, M. A. B., "The Electrical Activity of the Nervous System," 2nd ed., The Macmillan Co., New York, N. Y., 1960.

Cranefield, P. F. and Hoffman, B. F., "Electrophysiology of Single Cardiac Cells," *Physiol. Rev.* **38**, 41-76 (1958).

Eccles, J. C., "Neuron Physiology—Introduction," in "Handbook of Physiology," sec. 1, vol. 1, 59-74, 1959.

Keynes, R. D. and Martins-Ferreira, H., "Membrane Potentials in the Electroplates of the Electric Eel," *J. Physiol.* **119**, 315-351 (1953).

Lundberg, A., "Electrophysiology of Salivary Glands," *Physiol. Rev.* **38**, 21-40 (1958).

Tasaki, I., "Conduction of the Nerve Impulse," in "Handbook of Physiology," sec. 1, vol. 1, 75-122, 1959.

REPRODUCTION, GROWTH, AGING

REPRODUCTION is a basic characteristic of life. Underlying all forms of reproduction is cell division. In single cell organisms, of course, cell division results in reproduction of the entire organism. In multicellular organisms, cell division serves other purposes. In the first place, by the formation of additional cells the organism develops; secondly, such formation can replace cells that die; and finally, uncontrolled cell division and growth produces cancer. Quite clearly, a knowledge of cell division is essential.

MITOSIS

Mitosis is a complicated process in which the individual cell reproduces itself. Thus, at the end of the sequence of events, there are two cells and each one has the same number of chromosomes as had the original cell. A simpler form of cell division is termed **amitosis.** Again two cells result, but they do not have the same number of chromosomes. In other words, the cell has simply split so that part of the original cell is in one fragment and the remainder in the other. What role, if any, amitosis plays is uncertain.

Phases of Mitosis

Mitosis is a continuous process, but it has become the practice to divide it into four phases. It should be understood that the classic

description of the phases of mitosis does not apply to all cells. But in general, the dividing cell shows the following changes:

1. *Prophase*. Activity begins about the centrioles (Fig. 8.1). First, astral rays appear, then the centrioles move to the opposite poles of
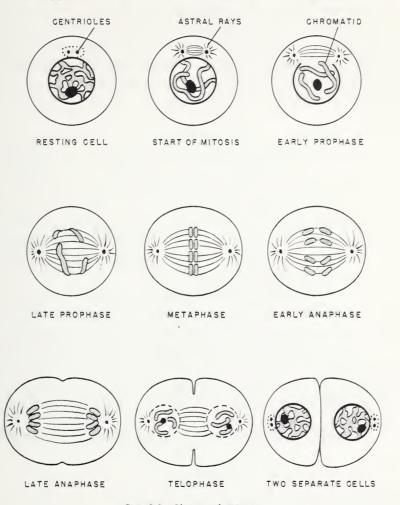
Fig. 8.1. Phases of Mitosis.

the cell. Meanwhile the chromosomes develop into distinct, separate entities; they become paired. Each member of the pair is termed a **chromatid**. At the end of the prophase the nuclear membrane disappears.

2. *Metaphase.* This is a short phase during which the chromosomes become aligned in an orderly fashion. Half of the chromatids face one centriole; half, the other.

3. *Anaphase.* In the anaphase the chromatids separate, half migrating toward one centriole; half, to the other.

4. *Telephase.* The chromatids reach the centrioles and the cell membrane forms a constriction. Then the cell splits into two distinct units. Meanwhile a nuclear membrane forms around each chromosomal mass. The centrioles return to their original positions and each new cell now begins to grow.

Duration of Mitosis

The duration of mitosis varies with the cell type, the temperature, and other factors. Accordingly, no specific figure can be given. From the published reports it would appear that it ranges from a few minutes to a few hours. Not only is the duration of mitosis highly variable, but so is the period of time between mitotic activity. Thus, a particular type of cell may take only a few minutes to undergo mitosis, but then it may remain in the resting or **interphase** for several hours before undergoing further division.

The duration of mitosis and the interval between mitotic activity can be altered in several ways. An increase in the environmental temperature speeds mitosis, as is to be expected. In an attempt to find an agent useful in the battle against cancer many substances have been shown to inhibit mitosis. Their value, however, in so far as cancer is concerned, still remains to be proved. And finally, there is a vast literature concerning the influence of radiation on mitosis. Various types of radiation do alter mitosis, but the effects on the cell are so widespread that more will be said concerning this subject in Chapter 9.

Initiation of Mitosis

No answer can yet be given to the question as to what initiates mitosis. The ovum does not divide, but rather quickly dies if it is not fertilized. In contradistinction, if the sperm enters the ovum, the ovum survives and undergoes rapid cell division. Why? Probably there is a chemical basis for the normal initiation of mitosis, yet it

has been shown that mitosis can be initiated by physical agents. For example, eggs have been caused to undergo division by the application of heat, or by being placed in hypertonic solutions, or by merely being pricked with a needle. Such a response is known as **artificial parthenogenesis.** Pathenogenesis means the development of an unfertilized egg.

Considerable work has been done using artificial parthenogenesis in an attempt to discover the initiating mechanism of mitosis. A change in the protoplasm seems to be an essential step. The protoplasm undergoes gelation, or clotting. This is said to be similar to the surface precipitation reaction (page 35) and in line with this hypothesis, calcium is found to be essential.

Not only does a change in the state of the protoplasm apparently play a role, but it would seem that the volume of the protoplasm is also important. There is considerable evidence to indicate that when the mass of protoplasm in a cell doubles, division occurs. Just how this triggers mitosis is not known. It has been suggested that by doubling the protoplasmic mass, the quantity of DNA would also be doubled and thus reach some threshold necessary for division. But, unfortunately, there is little evidence to support this hypothesis. Whether it is the simple increase in protoplasmic mass that accounts for the increase in DNA, or some other factor, is not known. Whatever the cause, the fact remains, that between cell divisions, the quantity of DNA in the cell does double.

Quite clearly the problem of cancer is intimately associated with the initiation of mitosis. Although it is possible that some means may be found that can be used clinically to inhibit cell division, certainly greater insight into the initiation of mitosis would help immeasurably in the search.

MEIOSIS

Individual cells reproduce by mitosis. But the reproduction of higher, multicellular organisms requires the fertilization of the ovum by the sperm. In mitosis each daughter cell ends up with exactly the same number of chromosomes possessed by the parent cell. In other words, during the prophase of mitosis the chromosomes appear as distinct, separate entities and then become paired, each

member of a pair an exact duplicate of the other. So there is a doubling and then half of the total number go to one cell and half to the other. In this way the chromosomal number is maintained. But in reproduction brought about by fertilization the sperm and the ovum each possess but half the normal number of chromosomes. When they come together the number is returned to normal. Thus, in the formation of sperm and ova a special kind of cell division is necessary in order to cut the chromosomal number in half. This type of cell division is termed meiosis.

Oogenesis

The formation of the ovum is termed oogenesis. In this process there is a halving of the number of chromosomes. There are generally two steps, or divisions. These have been termed **meiosis I** and **meiosis II.** It is in meiosis I that halving occurs. The chromosomal pairs line up and then one member of each pair goes to the two new cells. As shown in Fig. 8.2, the **primary oocyte** forms the **secondary oocyte** and the **first polar body** in meiosis I. The secondary oocyte and the first polar body possess but half the number of chromosomes of the primary oocyte. The secondary oocyte now undergoes division to produce the **ootid** and the **second polar body.** At the same time the first polar body also divides so that as a result of the two meiotic divisions there are four cells, each with half the chromosomal number. In this division there is no reduction of chromosomes, therefore meiosis II is sometimes referred to as meiotic mitosis. In short, the chromosomal halving occurs in the first division only.

Meiosis, however, is nowhere near as simple as just presented. It is an extremely complicated process, and there are important differences between mitosis and meiosis which should be appreciated. First, it must be understood that the higher organisms are **diploid.** This means that each cell possesses two similar sets of chromosomes. In mitosis each chromosome becomes duplicated so that there are identical pairs. Then one member of each pair goes to each daughter cell. Accordingly, each daughter cell has two similar sets of chromosomes, just as the parent cell did. In meiosis I the similar chromosomes, which are termed **homologues,** separate, one going to one cell, the other to another cell. Thus, each cell now has but half the original number of chromosomes; it is **haploid.** While this is hap-

pening each homologue becomes paired, but the pairs remain at-
tached at the centromere. In meiosis II these pairs separate, one
member going to one cell, the other to another. Thus, as stated
above, in the second division there is no further reduction of chromo-
somes; each cell is still a haploid.
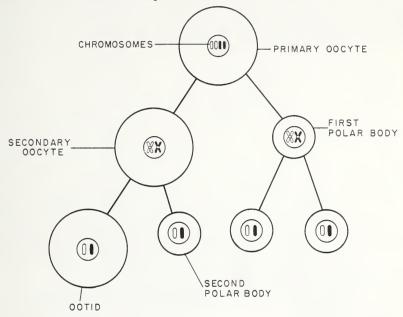
Fig. 8.2. Meiosis. Note that in the first division (meiosis I) the number of
chromosomes is cut in half. In the second division (meiosis II) the chromosomal
number is maintained.

Spermatogenesis

The process of spermatogenesis is very similar to oogenesis. In this
case, a **primary spermatocyte** divides, in meiosis I, into two **secondary
spermatocytes.** Each of these secondary spermatocytes has but half
the total number of chromosomes. Each spermatocyte then divides
again in meiosis II, but, as has been pointed out, the number of
chromosomes is maintained in each of the **spermatids.** Thus, from
each primary spermatocyte there arise four spermatids, each with half
the normal number of chromosomes. When the spermatid under-
goes further maturation and develops an elongated flagellum, it is
known as a **spermatozoon.**

CHROMOSOMAL FUNCTION

It has been stressed that in mitosis which represents reproduction of individual cells the number of chromosomes is maintained. These chromosomes are paired and, appropriately, are referred to as diploid. Reproduction of higher multicellular organisms requires fertilization of the ovum by a sperm. Each of these cells is in the haploid condition, that is, they contain but one member of each chromosomal pair. Such a cell is also known as a **gamete.** When the sperm and ovum come together the diploid condition is restored. This coming together to restore the chromosomal number is termed **syngamy.** The chromosomes, as will be discussed presently, play an essential role in reproduction.
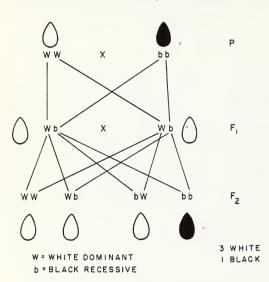
Fig. 8.3. Mendel's Law of Segregation. (See text for explanation)

Laws of Inheritance

Gregor Mendel made fundamental observations concerning patterns of inheritance. These observations were published in 1865 and have since been amply confirmed. He formulated two basic laws of inheritance:

1. *Law of Segregation.* Mendel assumed that there were paired genetic determinants. From his observations he concluded that these factors must segregate from each other in the parent and come together again in the offspring. For example, as shown in Fig. 8.3, assume that there is a pair of black characteristics in one parent and a pair of white characteristics in the other. In the first generation, F_1, both combinations are Wb; both white if white is **dominant.** In the next generation, F_2, according to the law of segregation, there

are four possibilities: 1) WW, 2) Wb, 3) bW, and 4) bb. There is a 3:1 ratio, 3 white and 1 black.

2. *Law of Independent Assortment.* Mendel also showed that when there is more than one pair of characteristics, or **alleles**, the pairs segregate from each other independently. Figure 8.4 shows that if there are two pairs there will be 16 possibilities in the F_2 generation. As will be discussed below, the second law does not always hold because of what is known as **linked genes.**

	WL	Ws	bL	bs
WL	WWLL	WWLs	WbLL	WbLs
Ws	WWLs	WWss	WbLs	Wbss
bL	WbLL	WbLs	bbLL	bbLs
bs	WbLs	Wbss	bbLs	bbss

9 WHITE LARGE
3 WHITE SMALL
3 BLACK LARGE
1 BLACK SMALL

W = WHITE DOMINANT   b = BLACK RECESSIVE
L = LARGE DOMINANT   s = SMALL RECESSIVE
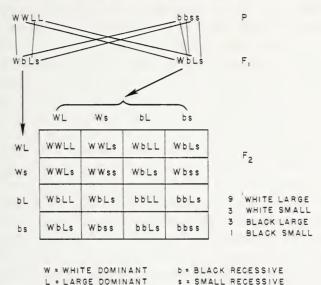
Fig. 8.4. Mendel's Law of Independent Assortment. (See text for explanation)

The amazing aspect of the Mendelian conclusions is that he had no knowledge of chromosomes. In other words, simply from his observations he was able to conclude that there must be specific factors in the cell that control the characteristics of the offspring in a predictable, orderly fashion. It is now known, of course, that the chromosomes have this function.

Genes

It early became quite obvious that each chromosome was responsible for many inheritable characteristics. Thus, the concept of the

gene arose in which a specific part of the chromosome is visualized as being responsible for one hereditary factor. There is now excellent evidence that the gene has a chemical nature. As mentioned when the nucleus was considered in Chapter 1, there is the growing belief that the storage and perpetuation of hereditary information, that is the gene, is deoxyribonucleic acid (DNA) located on a specific part of the chromosome.

The most fascinating current problem facing the cell physiologist is to elucidate how DNA functions as a transmittor of hereditary characteristics. Apparently the structure and metabolism of DNA can vary in a remarkable number of ways. And it would seem that each variation is capable of controlling a basic factor that results in the inheritance of a specific characteristic. Since variations in the structure and metabolism of DNA are thought to be the basic controlling factors, the concept of "coding" has arisen. In other words, DNA is looked upon as something akin to an IBM card which carries a code. Whenever the card is fed into an appropiate decoder, there is always a specific response. In view of the fact that the number of variations possible for DNA is tremendous all imaginable characteristics could be coded and transmitted in this way. It must be understood that this entire concept of DNA coding has got almost out of hand. Magnificent, but fanciful patterns for possible codes have been constructed so that one can quickly forget that they are based far more on imagination than experimental data. Thus, though truly incredible work has been done, there is cause for caution.

The hypothesis is growing that each gene is concerned with the biosynthesis of an individual protein. In this view, the resulting organism is but the summation of the physical and catalytic effects produced by a mass of specific protein molecules. In other words, the genetic information is somehow stored in the gene in the form of DNA. This gene, either alone, or perhaps in collaboration, with other genes is responsible for the production of a specific protein. The protein, then, in some manner gives rise to a particular characteristic. A hypothesis has been postulated which states that the genetic information within the DNA of the nucleus of the cell is transmitted to the nucleolus. Here the information is converted into RNA, and then the RNA passes into the cytoplasm to direct the

biosynthesis of cytoplasmic proteins. Whether DNA can participate directly in the synthesis of protein or always acts through an intermediate agent (RNA) cannot be decided on the basis of presently available evidence. Even if DNA can, under certain circumstances act directly, the currently available evidence strongly indicates that RNA has the major role in the mediation of nuclear-cytoplasmic interactions. It seems certain that the nucleus synthesizes RNA, but whether RNA can also be synthesized in the cytoplasm is not clear.

As mentioned in Chapter 1, the evidence is strong that genetic DNA occurs in the form of the so-called Watson-Crick double polynucleotide helices. This structure suggests that replication is accomplished by separation of paired nucleotide chains. Each chain is visualized as remaining intact to serve as a "template" against which new complementary partners may be built. These templates of DNA, that is genes, according to the current hypothesis, transfer specificity to RNA. The RNA made in this way migrates from the nucleus to the microsomes of the cytoplasm. In the microsomes, the RNA serves as a final template to direct protein synthesis. By this sequence a linear gene code is responsible for the linear arrangement of amino acids in a polypeptide chain.

As can be gathered from the above, the present-day thinking views each species of animal and plant as possessing specifically different proteins. And in each member of any one species there are more subtle protein differences to account for individual characteristics.

Linkage and Crossing Over

Exceptions to Mendel's second law early gave rise to the concept of linked genes. If the chromosome is thought of as having a series of genes, unless the chromosome itself splits into fragments, the genes of any one chromosome must move together, that is, they are linked and cannot segregate independently. But during meiosis there is separation of some of the genes of a single chromosome because there is an exchange of chromosomal segments between two chromosomal strands. This is termed crossing over. Quite clearly, the greater the distance between any two genes, the greater the possibility of independent segregation. On the other hand, the closer together they are the greater is the chance of them remaining linked.

Sex Determination

The chromosomal pattern has also been shown to be responsible for sex determination. In addition to the usual chromosomes, the so-called **autosomes,** there is also a special chromosome that is primarily concerned with sex determination. This is the X chromosome. In some species it has no partner, but in others its partner, the Y chromosome, carries no genes. In either case, as shown in Fig. 8.5 the ratio of male to female in the offspring must be 1:1.

Although the primary determination of sex is the function of the X and Y chromosomes, it is now known that genes located on the autosomes also play small, but nonetheless additive roles. In this light, sex determination depends on a ratio between two systems of genes each having small effects, some being localized on the X chromosome and the others on the autosomes. These genetic factors direct embryonic differentiation toward male or female development, but other influences can alter or dominate the action of these genetic factors. Thus, it is possible for a male to have the genetic make-up of a female, that is the XX combination, and for a female to possess the XY combination. The implications of this finding go beyond sex determination. It becomes clear that the genetic pattern through its coding system initiates mechanisms that can then be strikingly altered by other influences, at least during embryonic life, and probably later as well.
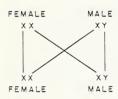
FEMALE   MALE
XX   XY

XX   XY
FEMALE   MALE

Fig. 8.5. Sex determination.

Sex-Linkage

The X chromosome carries many genes and unless there is crossing-over, the characteristics resulting from these genes must be sex-linked. Returning to the white and black example it can be seen in Fig. 8.6 that if the dominant white factor is on both X chromosomes in the female and the recessive black factor is on the X chromosome in the male then in the F_2 generation there is still the expected 3:1 ratio, but black can only appear in the male. There is no possibility of obtaining a black female.
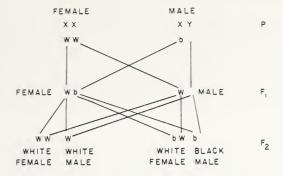
Fig. 8.6. Sex-linkage. The Y chromosome carries no genes. With the dominant characteristic on both X chromosomes in the female and the recessive in the male, the ratio in the F_2 generation is 3:1.

Another arrangement is shown in Fig. 8.7. Now the female is black because both X chromosomes carry the recessive gene. In mating with a white male there will be a 2:2 ratio in the F_2 generation with a white male and a white female, and a black male and a black female.

The results obtained in both examples depend upon a sex-linked gene plus a Y chromosome that is devoid of genes.
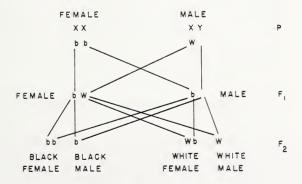
Fig. 8.7. Sex-linkage. The Y chromosome carries no genes. When the recessive characteristic is on both X chromosomes in the female and the dominant is on the X in the male, the ratio is 2:2.

GROWTH

The term growth means many things. It may imply enlargement of a cell or of an organism, or the term may have reference to the increase in size of a population. Basically, of course, growth means enlargement. This can be brought about either by the increase in size of individual cells, or by the multiplication of cells.

Growth Curve

It is quite interesting that growth, be it of a cell, an organism, or a population often describes an S-shaped curve (Fig. 8.8). It is seen
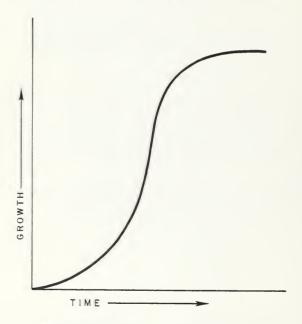
Fig. 8.8. Typical S-shaped Growth Curve.

that the increase in growth rate does so exponentially up to a point and then it levels off. The intriguing question is: why the limitation on size? One explanation is the increasing difficulty of obtaining adequate nourishment. Population studies show that the levelling off occurs when the population outruns its food supply. The same factor may play a role in limiting cell and organism growth. Insofar

as a cell is concerned, its supply of nutrients depends upon its surface area. As a cell grows the mass increases out of proportion to its surface area. This is true because in a sphere the **surface area** increases in proportion with the square of the diameter whereas the **total mass** increases with the cube of the diameter. Thus, just as populations cease to grow when they outrun their food supply, the same may be true for the individual cell. This, however, is undoubtedly but one factor, and probably an extreme one. Certainly, in some way, the genes, through specific protein synthesis regulate cell and organism growth. But just how, is not known.

Cellular Growth

Cells grow as a result of biosynthesis of protein, carbohydrate and lipid. Plant cells convert inorganic constituents by photosynthesis into carbohydrates. They also seem to be capable of synthesizing proteins and lipids. Thus, given an adequate supply of the basic inorganic materials and ample light, protoplasmic growth can proceed. The animal cell is also capable of biosynthesis but it depends ultimately upon plant cells for the basic building blocks, namely monosaccharides, amino acids, and fatty acids. From these, protoplasm is formed and the cell grows. As was discussed earlier, all of the reactions depend upon enzymes. The enzymes, which are proteins, are probably formed in accord with the DNA coding of the genes. It is, according to this view, the enzyme pattern which determines and regulates biosynthesis, thus controlling growth. For this reason mice do not become men regardless of the nutrient supply.

AGING

Why does protoplasmic growth stop? Why do cells age? And why do cells die? These are questions that, thus far, have not been satisfactorily answered. It was pointed out above that there is a rather characteristic growth curve that ultimately levels or plateaus. Does this plateauing represent aging, or is it simply due to limitations of food supply? In many instances food limitation is the cause, but this does not appear always to be the answer. There are several theories of aging. Most of them may be classified under two headings.

Toxicity

Cultures of microorganisms and other lower forms can be maintained seemingly indefinitely so long as the culture medium provides sufficient nutrition and is prevented from accumulating the waste products of metabolism, that is, toxins. When these conditions exist, organisms continue to grow and divide and thus appear to be immortal. This concept of toxicity as being the cause of aging is supported by experiments in which tissues from higher forms have been maintained in cultures for exceedingly long periods of time. Perhaps the most famous experiment of this type involved the maintenance of a piece of chicken heart in a culture for some 33 years, a period far longer than the normal life expectancy of the species. And, after 33 years the experiment was ended not because the tissue had aged and died, but for other reasons. All that was necessary to maintain the heart tissue that length of time was constant renewal of the medium so as to provide adequate nourishment and to remove metabolic toxins.

Protein Changes

Under normal conditions of aging (if aging can be said to be normal), there are definite changes in protoplasmic protein. In view of the fact that protein plays such an essential role in inheritance and in protoplasmic growth, it is not surprising to find that protein alterations also cause, or at least accompany, aging. Since it is the protein configuration that determines the individual characteristics, then there is a basis for the well-documented factor of inherited longevity.

Still, finding that there are protein changes hardly explains what causes them. Interestingly, it has been shown, again using tissue cultures, that if the medium is not changed the rate of growth and division decreases, and definite signs of aging develop, including characteristic protein alterations. Now if the aging cells are transferred to a new medium, after a lag they once again grow and divide and, significantly, the protein pattern is no longer that of old cells. Likewise, tissues have been taken from an old organism and placed in a culture. In this new environment they become rejuvenated, at least so far as their protein age is concerned. This does not prove that aging is caused simply by the accumulation of toxins, but it certainly suggests that toxins play an important role.

It should be understood that there is a constant turnover of protein. That is to say, protein is constantly being catabolized and new protein formed. It is thought that perhaps aging is associated with a decrease in the protein turn-over time; thus instead of new protein being formed at a vigorous rate, the same protein molecule in the aging cell persists and undergoes the alterations that have already been mentioned. In accord with this possibility is the observation that the ability of the cell to accumulate amino acids, the basic building units of protein, seems to decrease with age.

CANCER

Unlimited, uncontrolled cell growth and division is termed cancer. This is a malignant type of growth. Malignant growths not only invade surrounding tissues which are then destroyed, but clusters of malignant cells are carried by the blood and lymphatic circulation to other areas of the body where they give rise to secondary growths, called **metastases.**

Cancer Cell Morphology

There is a general tendency for cancer cells to become undifferentiated. Apparently the more rapid the growth, the more the cells resemble the embryonal form. This is not invariably true however. In some cases the cells retain the characteristics and functions of the tissues from which they arose. For example, a cancer involving a gland may secrete the same products as the normal gland in tremendous quantities.

Usually there is not only rapid cell division, but the individual cells grow abnormally large. The number of mitochondria increase as do the Golgi bodies. And, characteristic of a rapidly dividing cell, the nucleus is generally quite large. In some cases the nucleus is of normal size, but the nucleolus is abnormally large.

In a mass of cancerous cells there are usually many necrotic cells. Although there is accompanying growth of vascular and lymph channels, the cells seem to outrun the blood supply; thus there are nutrient deficiencies as well as the accumulation of toxins. This not only causes necrosis of some of the cancer cells, but the toxins may also give rise to deleterious effects throughout the organism.

Mitosis in Cancer Cells

There are characteristic mitotic abnormalities in malignant cells. The chromosomes vary in number and size. The separation of the chromatids may be delayed, thus resulting in unequal distribution of chromosomes in the daughter cells. In some instances the nucleus divides, but the cytoplasm does not, thereby giving rise to a multi-nuclear cell.

As might be expected, the quantity of DNA in the nucleus of the rapidly dividing cancerous cell is very great. But whether this is a cause of the unlimited, uncontrolled growth and division, or merely an accompanying characteristic is difficult to decide.

Etiology of Cancer

The etiology, that is, the cause, of cancer is a question that is not only of interest to clinical medicine, but of equal import to the cell physiologist. Why does a cell that has undergone normal, regulated growth, suddenly break away and embark upon unrestricted growth and division? There are, of course, several theories. It has been suggested that a virus invades the cells and somehow upsets the intracellular mechanism. On the other hand, the demonstration of a wide variety of carcinogenic agents, that is, substances that provoke cancer formation, gives rise to the concept that an abnormal environment in some way disrupts intracellular metabolism, specifically the respiratory mechanism. This is thought to liberate the cell from a limiting factor so that unrestricted growth with little cell differentiation occurs. And finally, with ever greater knowledge of the chemical basis of heredity and cell control, there are authorities who insist that the fundamental derangement involves the DNA-RNA mechanism. Quite clearly, greater knowledge concerning cellular physiology is a prerequisite to the understanding of cancer.

SUMMARY

In **mitosis** there is splitting of the chromosomes so that each daughter cell contains the same number of chromosomes as does the original cell. Cells may also simply split into fragments that are not necessarily identical. That is termed **amitosis.** Sperm and ovum are formed by **meiosis,** a process in which each cell receives but one mem-

ber of each chromosomal pair. An unfertilized egg can be caused to undergo division and development. This is termed **artificial parthenogenesis.**

The chromosomes carry genes which are probably **deoxyribonucleic acid** (DNA). The genes are responsible for the transmission of hereditary characteristics. DNA variations are thought to constitute a coding system which regulates protein synthesis and in this way determines all of the characteristics of the species. The two basic laws of inheritance are: 1) the law of segregation, and 2) the law of independent assortment. If the genes are linked the second law will not hold unless crossing-over occurs. The X and Y chromosomes are primarily concerned with sex determination, thus there is generally a 1:1 ratio of males to females. The **autosomes** also carry genes that can affect sex-determination. Influences operative during embryonic life can further alter sexual development. Thus an individual may have the genetic make-up of one sex, but have the physical characteristics of the opposite sex. The X chromosome also carries many genes that control characteristics other than sex. They are, therefore, sex-linked.

Very little is known concerning the mechanism of growth and aging. In tissue cultures, cells survive and continue to divide apparently indefinitely so long as there is ample food, and waste products do not accumulate. But whether aging of the cell or of the intact organism is due to toxins or to protein alterations is not known.

Unlimited, uncontrolled cell growth and division is termed **cancer.** The alterations that occur in cancerous cells have been well described but why cells should break away from the usual restraints to undergo unrestricted growth and division is, as yet, not known.

Problems

1. Explain the difference between mitosis, amitosis, and meiosis.
2. Assume that the male has two dominant red genes, the female two recessive black genes. In addition, the X chromosome in the male has a recessive short characteristic and both X chromosomes in the female have the dominant tall factor. Make a chart showing the possibilities in the F_1 and F_2 generations.
3. Outline the current concept of the chemical basis of heredity.
4. What is the evidence which supports the toxicity theory of aging?

5. Define the following terms:

a. Parthenogenesis	d. Metastases
b. Allele	e. Cancer
c. Syngamy	f. Oogenesis

Additional Reading

Anfinsen, C. B., "The Molecular Basis of Evolution," John Wiley & Sons, Inc., New York, N. Y., 1959.

Beadle, G. W., "Physiological Aspects of Genetics," *Ann. Rev. Physiol.,* **22,** 45-74 (1960).

Hughes, A., "The Mitotic Cycle," Scientific Publications, Butterworths, London, England, 1952.

McElroy, W. D. and Glass, B., eds., "A Symposium on the Chemical Basis of Heredity," The Johns Hopkins Press, Baltimore, Md., 1957.

Prescott, D. M., "Nuclear Function and Nuclear-Cytoplasmic Interactions," *Ann. Rev. Physiol.,* **22,** 17-44 (1960).

Swanson, C. P., "Cytology and Cytogenetics," Prentice-Hall, Inc., Englewood Cliffs, N. J., 1957.

Taylor, J. H., "Chromosome Reproduction and the Problem of Coding and Transmitting the Genetic Heritage," *Am. Scientist,* **48,** 365-382 (1960).

CHAPTER 9

EFFECTS OF
RADIATION

IT HAS ALREADY been mentioned that one type of radiation, that of visible light, can be used by the plant cell for photosynthesis. Light is also used by various organisms to regulate cycles or to initiate certain processes. But the radiations of visible light, in some instances, have a deleterious influence on cellular activities. This is certainly true of other types of radiations ranging from ultraviolet light to gamma radiations.

Radiations emitted from unstable atoms not only may have a profound effect on cellular activity, but they also provide the physiologist with one of his most powerful and versatile research tools. For this reason considerable space is devoted in this chapter to a discussion of atomic energy. In 1896 Henri Becquerel initiated interest in this field by discovering the radioactivity of uranium. Many investigators, such as Pierre and Marie Curie, quickly began the search for other sources of radiation. At the same time, Thompson, Rutherford, Bohr, and others devoted their efforts to a study of the atom itself. The true potential of atomic energy was made clear by the famous Einstein equation which states that if mass is converted to energy, the energy will be equal to the mass times the speed of light squared. In view of the fact that the speed of light is some 300,000 kilometers per second it is obvious that squaring this figure produces a very large number. A rough idea of what this means is obtained from the realization that a kilogram of any substance, if

completely converted to energy, would produce about 700 million horsepower!

The history of this field leading up to the atomic explosions of the second world war is a fascinating one. At the present time knowledge has progressed to the point where radioactive isotopes are readily available, and radioisotope equipment is commonplace in the research laboratory.

THE ATOM

The concept of the atom is very old. As early as the fourth century B. C., the Greek philosopher, Democritus, postulated that all matter is composed of very tiny particles, which he called **atoma.**

Anatomy of the Atom

According to the present atomic concept, each atom contains a central core termed the **nucleus.** Surrounding and moving in orbits around the nucleus are **electrons.** Figure 9.1 depicts a few typical atoms. From an illustration it is difficult to gain a concept of the relative mass of the nucleus and the electrons. The nucleus is so dense that it has been estimated that if a child's marble had the same density it would weigh many million tons!

Electrons have a negative charge. The nucleus is then believed to contain positively charged particles termed **protons.** In addition to the protons, the nucleus of the atom contains particles that are not charged and, accordingly, are called **neutrons.**

Atomic Number

The chemical property of a particular atom is apparently determined by the number of protons its nucleus contains. This number is termed the **atomic number.** For example, the hydrogen atom has but one proton, the oxygen atom 8, while uranium has 92. Usually, the number of electrons in orbit around the nucleus equals the atomic number, that is, the number of protons. However, it is possible to alter the electron count without changing the chemical characteristics of the atom.

Atomic Weight

The number of neutrons in the oxygen atom is the same as the number of protons. But this is not true of all atoms. Hydrogen, for
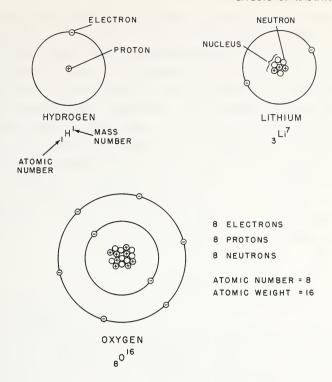
Fig. 9.1. Atomic Structure. The nucleus, except in the case of hydrogen, is made up of protons and neutrons. The electrons are in orbit about the nucleus.

example, has no neutrons while in the heavier atoms there are more neutrons than protons. The sum of the protons and neutrons is called the atomic weight. Thus for oxygen it is 16, while for lithium it would be 7, and for uranium 238. It was stated that the atomic number determines the chemical characteristics. The atomic weight indicates the physical properties. The atomic weight may be altered without influencing the chemical nature of the atom.

Isotopes

Students who are familiar with the Periodic Table will wonder why the atomic weights for the various elements are seldom whole numbers. For example, iron is 55.85 and copper is 65.54. The explanation is that elements are actually mixtures of somewhat different atoms. They are identical insofar as their chemical properties are

concerned, but they differ physically. Such atoms are called **isotopes.**
In other words, isotopes of the same element have different atomic
weights, more correctly, **mass numbers.** Closer study discloses that
the number of electrons does not change from one isotope to another.
For this reason the chemical properties are identical. And since the
protons generally balance the electrons, their number does not
change. The difference in atomic weight, then, must be due to an
altered number of neutrons. The conventional way to indicate the
atomic weight (mass number) of a particular isotope is to use a super-
script, for example, one isotope of potassium is written K^{42}, while
another is written K^{39}. For completeness the number of protons, that
is, the atomic number, is often indicated by using a subscript. For
example, potassium would then be $_{19}K^{42}$, or $_{19}K^{39}$.

RADIATION

It has been mentioned that electrons move, with great velocity, in
orbit about the nucleus of the atom. Why do they stay in orbit?
Electrons remain in orbit about the nucleus for the same reason that
man-made satellites remain in orbit about the earth. The "bucket-of-
water" analogy makes these relationships clear. As is well known, if
one swings a bucket of water in an arc with sufficient velocity, the
water will not spill out even though at the top of the arc the bucket
is upside down. The explanation is that the centrifugal force pro-
duced by the rotational velocity suffices to offset the force of gravity.
The electron is attracted by the nucleus and if the electron had no
velocity it would be pulled to the nucleus. However, it rotates with
sufficient velocity so that the centrifugal force balances the attraction
of the nucleus, and the electron then remains in orbit. This is true
of stable atoms; unstable atoms become altered by the emission of
radiation.

Types of Radiation

There are two general types of radiation: 1) **particulate** or **corpus-
cular** and 2) **electromagnetic.** In the first type of radiation there are
many possible types of particles. Among them may be mentioned
the alpha and beta particles, the neutron, proton, and positron. Only

the alpha and beta particle will be discussed here. The **alpha** particle has a weight of about 6.6×10^{-24} gram and has 2 positive charges. Each charge has been found to be equal to 4.8×10^{-10} electrostatic units. The **beta** particle is an electron which has a weight of about 9.2×10^{-28} gram and 1 negative charge. An alpha particle has the same composition as the nucleus of a helium atom (Fig. 9.2).

Electromagnetic radiation consists of small units of energy termed **photons** or **quanta**. The photon travels at the speed of light. In contrast, particulate radiation always travels at a speed less than that of light. Alpha particles move with a speed of about 20,000 kilometers per second. Beta particles approach the speed of light which is close to 300,000 kilometers per second.

HELIUM   ALPHA PARTICLE   BETA PARTICLE
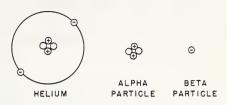
Fig. 9.2. Alpha and Beta Particles Shown in Relation to a Helium Atom. The alpha particle is identical in composition to a helium nucleus. The beta particle is an electron.

The energy and the frequency of electromagnetic waves are inversely proportional to the wavelength. Thus ultraviolet light has very little energy whereas short-wave X-ray and gamma radiation have very high energy (Fig. 9.3).

The wave length of the gamma radiation is about 10^{-10} centimeter. This is the shortest wavelength in the electromagnetic spectrum. At the other end are the radio waves which measure from 10^{-1} to 10^5 centimeters. In the middle of the spectrum is visible light. When an unstable nucleus emits gamma rays, the nucleus is not markedly changed as it is when alpha or beta particles are emitted. The major change is that the excited nucleus reduces its excess energy by the emission of gamma radiation.

Pattern of Disintegration

When an unstable atom emits radiation the atom changes in either structure or energy level, or both. This change is termed disintegration which is not really an accurate term to describe the process since there is certainly not complete decomposition or destruction. However, the term has won wide acceptance to describe the alterations that occur in unstable atoms.
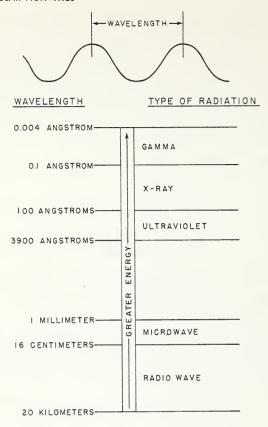
Fig. 9.3. The Wavelength of Various Types of Radiations. The shorter the wavelength the greater the energy of radiation.

A specific atom always disintegrates in the same way and at the same rate. Thus if a particular unstable atom disintegrates by emitting an alpha particle it will always do so. Likewise, some atoms emit beta particles, or gamma rays. There are atoms that give off all three types, but always in a specific pattern.

In view of the fact that an alpha particle contains two protons and two neutrons, and a beta particle is an electron, it follows that the emission of these particles changes an atom of one element into an atom of an entirely different element. For example, when uranium disintegrates an alpha particle is given off. Thus, $_{92}U^{238}$ becomes

$_{90}\text{Th}^{234}$. This new element, thorium, is also unstable. It disintegrates by emitting a beta particle, that is, an electron. Thus the atomic weight does not alter, but with the loss of an electron a neutron is converted into a proton and, accordingly, the atomic number increases from 90 to 91 which is the element protoactinium. Protoactinium also emits a beta particle and so, once again uranium results, but now with an atomic weight of 234, that is, an isotope of the original uranium-238. Uranium-234 is also unstable and emits an alpha particle. In this way, a whole series of elements is produced before a stable atom of $_{82}\text{Pb}^{206}$ results.

Rate of Disintegration

It was just stated that not only does an unstable atom disintegrate following a specific pattern, but also at a set rate. To put it another way, a fixed percentage of the atoms present disintegrate per unit time. The interesting aspect of this phenomenon is that, unlike a chemical reaction, such factors as temperature and pressure exert no influence on the rate of disintegration.
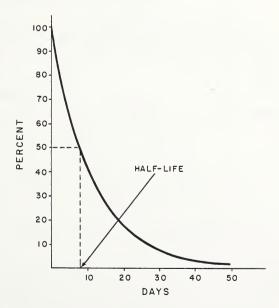
Fig. 9.4.   Decay Curve for I^{131}. The time required for 50 percent decay is the half-life.

The rate of disintegration is most commonly expressed in terms of **half-life.** The half-life is the period of time in which half of all the atoms of an unstable element will disintegrate. Again, let it be emphasized that by disintegrate is meant not complete destruction, but rather merely a change due to the emission of radiation.

Figure 9.4 depicts a time curve for the disintegration of I^{131}. In view of the fact that half the existing atoms always disintegrate in a certain time an exponential curve results. Thus, in one half-life 50 percent of the atoms disintegrate, in two half-life periods 75 percent, and in 3 half-life periods 87.5 percent. Because an exponential curve appears as a straight line when plotted on semi-logarithmic graph paper, this is the usual procedure for the calculation of half-life (Fig. 9.5). It should be noted that after 7 half-life periods there is less than one percent of the original number of atoms remaining.
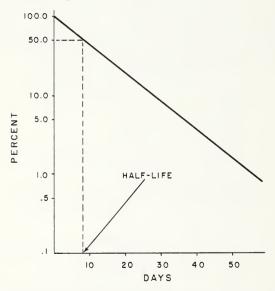
Fig. 9.5. Decay Curve for I^{131} Plotted on Semi-log Paper. A straight line results.

The half-life varies tremendously from one element to another. Table 9.1 lists but a few unstable elements to illustrate this point.

It is also worthy of note that isotopes of the same element may have a considerably different half-life. For example, for Na^{24} it is 15 hours while for Na^{22} it is 2.6 years.

TABLE 9.1. Half-Lives of Various Elements

Element	Half-Life
Potassium[42]	12.5 hours
Sodium[24]	15.0 hours
Mercury[197]	2.7 days
Iodine[131]	8.1 days
Iron[55]	2.9 years
Radium[226]	1620.0 years
Uranium[238]	4.6 billion years

Curie

The curie is the unit of activity of radioactive atoms. It has been shown that in 1 gram of radium, 3.71×10^{10} disintegrations occur per second. Therefore, any source having this same number of disintegrations per second is said to have an activity of 1 curie. This is a very large unit and for amounts of radioisotopes generally used in the biological research laboratory the millicurie, or even the microcurie usually suffices. Thus:

$$1 \text{ curie} = 3.71 \times 10^{10} \text{ disintegrations per second}$$
$$1 \text{ millicurie} = 3.71 \times 10^{7} \text{ disintegrations per second}$$
$$1 \text{ microcurie} = 3.71 \times 10^{4} \text{ disintegrations per second}$$

Energy of Radiation

The energy of each type of radiation can be determined. Energy can be expressed in ergs, or, as is more common, in terms of electron volts. Thus:

$$1 \text{ erg} = 6.24 \times 10^{5} \text{ million electron volts (Mev)}$$

The energy of particulate radiation depends upon the mass and velocity of the particle. Since alpha particles are relatively large, that is, they have a mass of 6.65×10^{-24} gram and travel at an average velocity of about 2×10^{9} centimeters per second, it may be calculated that the energy of such a particle is 1.3×10^{-5} erg or about 8 Mev. Depending on the velocity, the energy range of alpha particles is from 4 to 10 Mev.

The beta particle is smaller than the alpha. Its mass, as already mentioned, is 9.19×10^{-28} gram. But the velocity is greater, approaching that of light. The energy range of beta particles, again

depending upon the velocity, varies from about 0.2 to 3.8 Mev. It should be understood at this point that even though a beta particle may have less energy than an alpha particle it will penetrate matter to a much greater distance. This is due to the fact that the beta particle is much smaller than the alpha particle and it has but half the charge. This factor of penetration will be discussed more fully below.

Gamma rays have very short wavelengths and extremely high frequencies. They have the highest energy of all the components of the electromagnetic spectrum. The calculation of the energy of this type of radiation is rather complex. If one takes into consideration the frequency of the radiation, that is, the number of waves passing a given point per second, and a constant known as Planck's constant which is expressed in erg-seconds, the energy of a particular gamma radiation can be calculated. It is found that for most elements emitting this type of radiation the energy range is from about 0.1 to 2.75 Mev. Gamma radiation, however, has extreme power of penetration and thus constitutes a serious radiation hazard.

Absorption

Radiation from an unstable atom interacts with other atoms. The assimilation within an atom, or molecule, of energy from electromagnetic or particulate radiation is termed absorption. Thus, if the radiation strikes a thick layer of material, all of its energy may be completely dissipated due to interactions with the atoms of that material. In other words, the radiation will be completely absorbed. This is the basis of **shielding.**

The absorption for a particular radiation is determined by placing increasingly thick layers of materials such as aluminum or lead in the path of the radiation and then ascertaining the radiation that is transmitted through the material. Aluminum may be used for absorption studies of alpha and beta particles. The aluminum thickness necessary to absorb the alpha or beta radiation completely is determined. Thickness is usually expressed in grams per square centimeter. A graph is then used to determine the energy value corresponding with the aluminum thickness. In this manner the energy of radiation of an unknown isotope may be determined.

It is extremely difficult to absorb gamma radiation completely.

Therefore, it has been found more satisfactory to determine the **half-value thickness** of lead. In practice, one determines the radiation of the element under study first without any absorber in place. Then lead absorbers of increasing thickness are used and the radiation transmitted determined for each one. When the values are plotted, it is then possible to determine the thickness of lead required to cut the transmitted radiation in half. Again the graph may be consulted and from it the energy value of the gamma radiation ascertained.

BIOLOGICAL EFFECTS

Radioisotopes provide a superb tool for the investigator, but unfortunately they also constitute a serious health hazard. The scientists to whom we owe our first knowledge of radiation discovered this lamentable fact, in many cases, too late. Fortunately, increasing understanding of radiation has made it possible to use radioactive materials with complete safety.

Ionization

High energy radiation is injurious because as it penetrates matter, that is, the body, it interacts with the atoms which compose the body tissues. Interaction means disruption, a disturbance of organization and balance. Such disruption may involve a collision with an orbital electron of sufficient force to send it out of orbit. The electrical balance of the atom is thus upset and due to the loss of an electron the remaining atom is now positively charged. The ejected electron soon joins another atom which now has more electrons than protons and thus is negatively charged. This process by which neutral atoms become charged is termed ionization. Although the ejected electron may remain in the free state for some time, usually it ultimately joins another atom. Thus, in the ionization process two ions normally are formed; these constitute an **ion pair.**

Ionization may occur in other ways. For example, high velocity neutrons do not influence orbital electrons but they are capable of exciting the nucleus of an atom. The nucleus then may emit a proton which strikes an orbital electron and causes ionization. A continuing, or chain, reaction may occur in which radiation liberates an orbital

electron which then collides with an electron of another atom and dislodges it, etc.

Measurement of Ionizing Radiation

Several different units are used to state a quantity of radiation. They are all based on the ionization produced by the passage of the radiation through a medium. The units most likely to be encountered in radioisotope work are the roentgen, rep, rad, and rem.

The **roentgen** is a unit of radiation named in honor of the German physicist, Wilhelm Roentgen, who discovered X-rays. The roentgen is defined as the quantity of X- or gamma-radiation that will produce 1 electrostatic unit of charge, either negative or positive, in 1 cc of air at standard temperature and pressure, that is, 0°C and 760 mm Hg. It takes 2.08×10^9 ion pairs to provide one electrostatic unit of charge. The roentgen, therefore, is that quantity of X- or gamma-radiation that produces 2.08×10^9 ion pairs when it penetrates 1 cc of air. The production of this number of ion pairs represents an energy of 83 ergs per gram of dry air.

The **rep** is an abbreviation which stands for "roentgen equivalent, physical." It is a unit of radiation that expresses the ionization in soft tissue. The absorption of energy in tissue depends upon two factors: 1) the composition of the tissue, and 2) the energy of the radiation. The rep, therefore, is a unit which varies with the tissue and thus only an average value can be stated. One rep is generally considered to represent 83-93 ergs per gram of soft tissue for the dissipation of particulate radiation.

The **rad** is an abbreviation which stands for "radiation absorbed dose." It is a unit designed to overcome the objections to the rep which varies with the exposed tissue. By definition, one rad represents an energy absorption of 100 ergs per gram of any medium. This unit can, therefore, be applied to any type of radiation.

The **rem** is an abbreviation which stands for "roentgen equivalent, man." From the definitions of the roentgen, rep and rad, it should be clear that these units could not be used to evaluate easily the biological effects of all types of radiation on man. The rem was designed for this purpose. It is defined as the quantity of any type of radiation which produces the same biological effects in man as those resulting from the absorption of 1 roentgen of X- or gamma-radiation. This

biological effect varies with many factors, therefore, to calculate the rem one must use a **relative biological effectiveness** factor (RBE). Thus tables have been prepared listing the RBE for each biological effect and for each type of ionizing radiation. Since RBE values vary with so many factors such as distribution of the dose and the age of the individual, RBE values represent merely careful approximations.

Lethal Levels of Radiation

Radiation in sufficient quantity is lethal. Exposure to a quantity of radiation in excess of 500 rads to the entire body will prove fatal to over half of the individuals exposed. The sequence of events preceding death varies somewhat in different individuals but it usually involves nausea, diarrhea, fever, hemorrhage and delirium. If the victim lingers a week or more, profound anemia and a very low white blood cell count usually develop. And should he somehow survive for longer periods there may then develop leukemia or other forms of cancer. Thus, if such a large quantity of radiation does not prove immediately fatal, it will undoubtedly produce serious biological effects in time.

Effects on Cellular Mechanisms

It has been shown that a dosage of about 100 roentgens damages from 1 to 5 molecules of DNA in the mammalian cell. In view of the known roles of DNA, many of the long term effects of radiation become explicable. It is surmised, but not demonstrated, that radiation also alters RNA.

Radiation influences many biochemical mechanisms. For example, aerobic phosphorylation shows impairment within 30 minutes after irradiation by 50 roentgens.

The most conspicuous damage resulting from irradiation occurs in the chromosomes. The chromosomes undergo breaking in some areas, whereas in others they may adhere to one another and thus interfere with the normal process of mitosis. The influence of radiation on the chromosomes accounts for the well-documented instances of genetic mutations that so often result.

In view of the damage done by irradiation to the chromosomes it is not surprising that irradiation impairs mitosis. The effects are most profound if irradiation occurs during the phase of mitosis in which

the cell is devoid of both the nuclear membrane and the nucleolus. At this time comparatively small doses of radiation will delay mitosis; large doses completely stop it. In either case the cell usually dies, if not immediately, then after one or two divisions. The mechanism by which irradiation inhibits mitosis is not known. It has been postulated that it is somehow related to an influence of irradiation on DNA synthesis.

Irradiation causes cytoplasmic swelling and the development of giant cells. This, however, is not to be interpreted as a stimulus to growth. The giant cell represents impairment of division. Thus, the irradiated cell, instead of dividing, simply continues to enlarge and ultimately dies.

In some cells, irradiation causes swelling and other changes in the mitochondria. Along with these alterations, inhibition of respiration and of phosphorylation, an increase in ATP and an altered lipid metabolism have been reported. But whether these end results are due to altered mitochondrial function cannot be decided until there is more complete understanding of the role of mitochondria in the normal cell.

Microsomes are thought to be primarily concerned with protein synthesis. In this light it is significant that after irradiation protein synthesis impairment does not occur immediately. It would be interesting to know if irradiation damages the microsomes, and if so, whether the damage is immediate or delayed. Such experiments apparently have not yet been carried out.

Since irradiation impairs so many intracellular activities, it is not surprising that large doses decrease cell mobility. Thus, sperm become immobile as well as infertile. Also, phagocytosis has been reported to be impaired following irradiation.

Radiation is said to decrease the permeability of cell membranes. By this it is meant that following irradiation various substances are not transported by the cell membrane as readily as they are before exposure. But whether this result is due to a change in permeability or to an alteration in the specific transport mechanism cannot be said. Published reports show a decrease in intracellular potassium and a gain in sodium. It has also been shown that the entry of glucose and amino acids into irradiated cells is impaired.

Health Protection

In view of the fact that radiation may have such dire biological effects, permissible levels of exposures have been established. The unit of radiation used to express these levels is the rem. The basic permissible level depends upon the duration of exposure. An individual may receive a relatively large quantity of radiation for a very short period of time without ill effects. On the other hand, if he works with radioisotopes he may be exposed several hours a day for many years. Limits, then, have been placed on such long-term exposures. They have been set as follows:

1. For each 3-month period, a total of 3 rem
2. For each year, a total of 15 rem
3. For each 10-year period, a total of 50 rem

These limits mean that should a worker be exposed to relatively high quantities of radiation for a short period of time, say 3 rem, then he should not be further exposed, at all, for 3 months. To get these values down to a usuable range it has been estimated that an exposure of about 300 millirems per 40 hour week is permissible. This means an exposure of about 7-8 millirems per hour which is a good figure to remember since various instruments used to evaluate radiation are calibrated on an hour basis.

Shielding

In any radioisotope laboratory variable quantities of unstable atoms must be stored. In order to protect the personnel working in the area the radioisotopes are stored in such a way that the resulting radiation in the working area is below maximal permissible levels. It has been emphasized that radiation is absorbed by various materials. Alpha particles are easily and completely absorbed by most any material, even glass, or cardboard. Beta particles are more penetrating, but it is the gamma radiation that poses the major difficulty. For this purpose thick layers of lead are used for absorption, that is, to shield the isotopes. In the average biological laboratory using such unstable atoms, a small enclosure made of lead bricks generally suffices. The amount of lead shielding needed for storage of specific quantities of any isotope can be calculated, but in practice one simply places the material in the enclosure and then measures the radiation

outside the enclosure. The maximal permissible level is considered to be about 3 milliroentgen per hour.

SUMMARY

The **atom** is the basic structure of all matter. A substance with but one kind of atom is called an **element**. Each atom contains a central core termed the **nucleus.** Moving in orbits around the nucleus are **electrons**—particles that have a negative charge. Within the nucleus are protons with a positive charge and **neutrons** which are not charged. The number of protons determines the **atomic number.** The sum of the protons and neutrons is called the **atomic weight.** Isotopes have the same atomic number but different atomic weights. Isotopes of the same element have identical chemical properties; the difference in atomic weight is due to an altered number of neutrons.

An unstable atom that emits **radiation** is said to be **radioactive.** There are two general types of radiation: 1) particulate and 2) electromagnetic. The **alpha** particle is identical to the nucleus of a helium atom and is relatively large; .it has 2 positive charges. The **beta** particle is smaller and has one negative charge. The energy of alpha and beta particles is determined by their mass and velocity. **Gamma** rays are a form of electromagnetic radiation. They travel at the speed of light. The energy of electromagnetic waves is inversely proportional to the wavelength. Gamma rays have extremely short wavelengths and thus have very high energy.

The process of emitting radiation is termed **atomic disintegration.** The **half-life** is the period of time in which half of all the atoms of an unstable element disintegrate. The **curie** is the unit of activity of radioactive atoms. One curie equals 3.71×10^{10} disintegrations per second. The assimilation within an atom, or molecule, of energy from electromagnetic or particulate radiation is termed **absorption.**

High energy radiation is injurious because it penetrates the body and produces **ionization.** Units that express the quantity of radiation are based on the ionization produced by the passage of the radiation through a medium. The **roentgen** is defined in terms of ionization in air. The **rep** expresses ionization in soft tissue. The **rad** represents an energy absorption of 100 ergs per gram of any medium. The **rem** takes into consideration the biological effect.

Whole body irradiation in excess of 500 rads is usually fatal; in smaller doses there are widespread alterations in cellular activities. These include impairment of DNA and aerobic phosphorylation, chromosomal damage, prevention of mitosis, mitochondrial swelling, deranged protein synthesis, decreased cell motility, and diminished transport through the cell membrane.
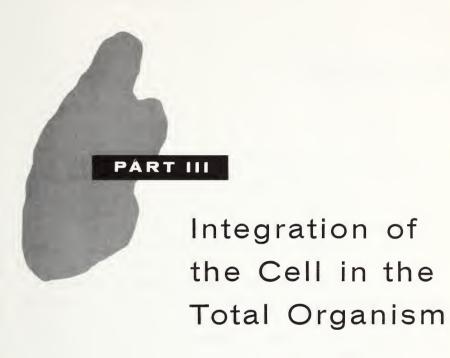
Problems

1. Make a drawing depicting the structure of an oxygen atom.
2. Explain the difference between atomic number and atomic weight.
3. An alpha particle may have the same energy as a beta particle. Which would be more penetrating? Why?
4. How may the energy of alpha, beta, and gamma radiation be determined?
5. Explain the difference between the roentgen, the rem, the rad, and the rep.
6. Assuming a chemical basis of heredity, explain how irradiation produces mutations.

Additional Reading

Browdsted, H. V., "The Atomic Age and Our Biological Future," Watts, London, England, 1957.

Claus, W. D., ed., "Radiation Biology and Medicine," Addison-Wesley Publishing Co., Reading, Mass., 1958.

Elton, C. R. B., "Introductory Nuclear Theory," Interscience, New York, N. Y., 1959.

Fermi, L., "Atoms in the Family," Chicago University Press, Chicago, Ill., 1954.

Kamen, M., "Radioactive Tracers in Biology," 3rd ed., Academic Press, New York, N. Y., 1957.

Kaplan, H. S., "Cellular Effects of Ionizing Radiation," *Bull. N. Y. Acad. Med.* **36**, 649-661 (1960).

Quimby, E. H., Feitelberg, S. and Silver, S., "Radioactive Isotopes in Clinical Practice," Lea & Febiger, Philadelphia, Pa., 1958.

PART III

Integration of
the Cell in the
Total Organism

CHAPTER 10

TISSUES

A GROUP OF similar cells united into a mass or structure is termed a tissue. The very same cells may arrange themselves in one way and yield a specific tissue. Another type of cell aggregate may make up a different kind of tissue. In short, because of the variety of cell types and the manner in which they are grouped, the wide selection of tissues characteristic of the more complex organisms is made available.

Tissues are classified in several ways. For example, they may be grouped on the basis of the origin of the cells. In this classification there are ectodermal, mesodermal, and entodermal tissues. It is usually more convenient, however, to consider their derivatives, that is, epithelial, connective, muscular, and nervous tissues. This is the classification that will be followed here.

TISSUE DEVELOPMENT

In man, as well as many other species, development begins with the fertilization of the ovum by the sperm. The ovum then begins to divide, thus forming a multicellular mass.

The first new cells formed by divisions of the ovum show very little, if any, differentiation. Ultimately, however, three distinct layers of cells result which can be differentiated. These basic embryonic layers are termed **germ layers.** They are named according to their position in the embryo: **ectoderm, mesoderm,** and **entoderm.**

The three germ layers undergo extensive development so that ultimately specialized tissues result. Tissues from a specific germ layer

may form into an organ, or it is possible for tissues from different germ layers to become associated in the formation of an organ. Yet, despite this association, or integration, it is generally still possible to determine from which germ layer each cell arose. Table 10.1 lists the origin of various tissues from the three germ layers.

EPITHELIUM

Epithelial tissue covers the surface of the body, and it also lines the body cavities. This type of tissue consists of one or more layers of cells with scarcely any intercellular substance so that the cells form a practically unbroken sheet or membrane. As a consequence of their position, it is largely through the epithelial cells that the organism comes into relationship with the external world. They serve, as in the epidermis, or skin, to enclose and protect the body. Through their substance, and largely by their activity, the body absorbs its

TABLE 10.1. Germ Layer Derivatives

Ectoderm	Mesoderm	Entoderm
Skin epithelium	Connective tissue	Respiratory epithelium
Mammary glands	Muscle	Bladder epithelium
Sweat glands	Blood vessel endothelium	GI tract epithelium
Nasal epithelium	Lymphatic endothelium	Thyroid gland
Mouth epithelium	Serous membrane mesothelium	Parathyroid gland
Anal epithelium	Urinary epithelium	Pancreas epithelium
Urethral epithelium	Adrenal cortex	Liver epithelium
Nerve		Gall bladder epithelium
Adrenal medulla		
Hypophyseal gland		

nutrient and excretes its waste products; they are also the active secreting elements of glands.

Types of Epithelial Tissues

Epithelial tissues are generally classified according to the shape of the component cells which vary from the very flat type which is termed **squamous,** to the tall cells called **columnar.** Between these two extremes are the **cuboidal** cells. Thus, there is squamous, cuboidal, and columnar epithelium. In all three, the tissue is referred

to as **simple** if the cells are arranged in a single layer. The tissue is said to be **stratified** if there is more than one layer of cells. Accordingly, there may be simple or stratified squamous epithelium, or simple columnar epithelium, and so forth.

Figure 10.1 shows various types of epithelia. The relatively flat cells of the squamous epithelium are arranged irregularly in a mosaic pattern. The nuclei are generally oval. Squamous epithelium lines

SQUAMOUS

CUBOIDAL

COLUMNAR

TRANSITIONAL
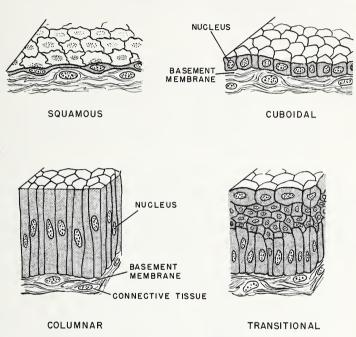
Fig. 10.1. Types of Epithelial Tissue.

the inner surface of the lungs, and is also found in the lens of the eye and in the inner ear. Stratified squamous epithelium forms the external layer of the skin and lines the mouth, pharynx, esophagus, vagina, and anus.

The cells of cuboidal epithelium form a very regular pattern. The nucleus is round and centrally placed. This type of epithelium is found lining small ducts and tubules.

There are many variations of columnar epithelium, but in all, the cells are taller than they are wide. The nucleus is oval-shaped and generally placed close to the bottom of the cell. Some of these cells

possess cilia. Ciliated epithelium lines the trachea and bronchi. Simple columnar epithelium is widely distributed in the digestive tract where it functions to secrete fluids and to absorb digested foodstuffs.

There is also a type of tissue which has many of the characteristics of both stratified squamous and columnar epithelial. For this reason it is termed **transitional epithelium.** This type of epithelium lines portions of the urinary tract.

Glands

A gland is an organ which secretes a fluid, generally of specific composition. The epithelial tissue of glands appears to be primarily responsible for the elaboration of the fluid secreted.

Broadly speaking, glands may be grouped as either exocrine or endocrine. An **exocrine** gland is one which secretes into a duct. The duct then carries the secretion to a specific part of the body, generally quite close to the gland. An **endocrine** gland secretes directly into the blood stream. For this reason, endocrine glands are also known as **ductless** glands.

The configuration of glands varies (Fig. 10.2). Simple glands (such as sweat glands) have a single unbranched duct. Compound glands
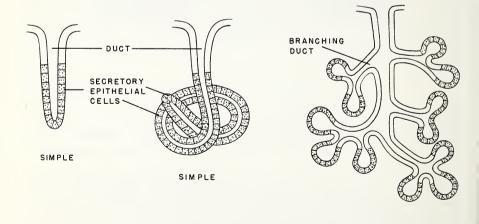
Fig. 10.2. Simple and Compound Glands. They differ primarily in that the duct branches in the compound gland. Note that the secretory cells are epithelium.

have a branching duct system. Compound glands are located in the liver, kidney, and testes. In all cases, the secreting layer is composed of epithelial cells.

Membranes

As has already been indicated, epithelial tissue forms important body membranes. These include **serous membranes** that line the pleural, pericardial, and peritoneal cavities and **mucous** membranes that line all cavities and canals of the body which connect with the outside. Thus, mucous membranes are found in the alimentary, respiratory and genito-urinary tracts.

The epithelial cells form only one part of the membrane; whereas the other part is composed of connective tissue. The underlying connective tissue thickens and toughens the membrane. It permits some stretch, but prevents excessive stretch which would separate the epithelial cells.

Epithelial Regeneration

If epithelium is damaged, it undergoes repair so that ultimately the injured cells are replaced by new ones. In general, when the surface cells are damaged, the deeper layers enlarge and migrate to the surface. These cells appear to move in a manner very similar to ameboid movement. They undergo rapid mitotic division so that the injured surface ultimately becomes covered with new cells.

CONNECTIVE TISSUE

Connective tissue, as the term indicates, serves as a binding substance. It forms the framework of the various organs and connects them to other structures throughout the body. This type of tissue is easy to distinguish from epithelium because of the great abundance of **intercellular** or ground substance. It was noted above that epithelium consists of rows of cells placed one against the other with little or no substance in between. Figure 10.3 shows that in connective tissue the intercellular material dominates the entire structure. The cells are few and scattered. As a matter of fact, it is the character of the intercellular substance which is used to classify connective tissue.

Areolar Connective Tissue

All connective tissue consists of cells, fibers, and a ground substance—a composition that makes the tissue soft and displaceable.

Areolar connective tissue is the most common type. The term areolar is derived from a word that means "space," and thus describes the characteristic spaces between the interlacing network of elastic fibers which make up the intercellular material. This type of tissue is widely distributed around organs, and it serves to protect and support them. Thus, if the organ is pulled out of its normal position, the strands of areolar tissue can be seen clinging to it.
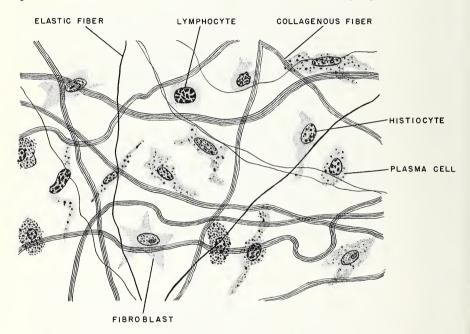
Fig. 10.3. Areolar Tissue.

The spaces in areolar tissue permit easy accumulation of fluid. When the quantity of fluid in these spaces becomes excessive, a condition exists which is termed **edema.**

Areolar tissue contains several cell types (Fig. 10.3). These include **fibroblasts, histiocytes, mast cells, plasma cells,** and cells from the blood, sometimes referred to as **wandering cells.** The fibroblasts are

most numerous and are responsible for the formation of fibers, thus their name. Following injury they form new fibers. The histiocytes, also known as **macrophages,** are the second most numerous cell type in areolar tissue; they are capable of ameboid movement and are phagocytic. Plasma cells are only rarely seen. They are thought to be derived from lymphocytes, but their function is obscure. The mast cells are thought to be responsible for the production of an anticoagulant, heparin. The cells from the blood include lymphocytes, eosinophils and neutrophilic leucocytes.

Areolar tissue contains three types of fibers. These are: 1) **white** or **collagenous fibers,** 2) **reticular fibers** and 3) **elastic fibers.**

It has been mentioned that because of the loose structure of areolar tissue, it contains a variable amount of fluid, called **tissue** or **interstitial fluid.** In addition there is the ground substance which is a glycoprotein. The ground substance is thought to play an important role in preventing the spread of infectious processes.

Areolar tissue, then, not only protects and supports various organs but also provides an anticoagulant to assist in healing and to limit and combat infection.

Fibrous Connective Tissue

Areolar tissue is also referred to as loose connective tissue. In contradistinction, fibrous tissue is known as dense connective tissue, because of the density of its fibers. The fibers are extremely strong and are generally quite wide. They cause ligaments, tendons, and fasciae to have a glistening white appearance.

Adipose Tissue

Fat accumulates in certain cells; in many areas of the body these cells are numerous and loosely organized into a mass termed adipose tissue. It is the cell and not the intercellular substance which makes up the bulk of this type of connective tissue (Fig. 10.4). The accumulation of the fat droplet in the cell displaces the nucleus to one side. Therefore if the fat is dissolved out of the cell, what remains appears empty and has a "signet ring" configuration. The fat in these cells is constantly undergoing catabolism. But if the caloric

intake equals energy needs, then the new formation of fat keeps pace with the rate of catabolism so that the mass of fat remains constant but is continuously changing. It is for this reason that if the caloric intake is reduced below body energy needs, the fat stores will decrease.

Adipose tissue not only represents an important reservoir of nutrient, but it also serves for the conservation of body heat and for protection against mechanical trauma.
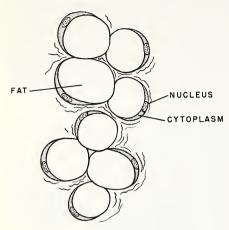
Fig. 10.4. Adipose Tissue. Note how the accumulation of fat within the cell forces the usual cell constituents to one side.

Cartilage

The ground substance in cartilage causes this type of connective tissue to be compact, slightly elastic, and remarkably strong. In the adult organism, cartilage covers the ends of the long bones, thus forming the articular surfaces. It also supports and protects many structures such as the nose, larynx, trachea, and bronchi.

Three types of cartilage are recognized: 1) **hyaline,** 2) **elastic,** and 3) **fibrous.** Hyaline cartilage is the most common type. It appears as a blue-white, translucent mass in the living state. The cells are widely separated by the ground substance, and occupy spaces in the ground substance known as **lacunae** (Fig. 10.5). In the ground substance there are very fine collagenous fibers. Elastic cartilage is more yellow and opaque than hyaline cartilage, due to the elastic fibers which form a dense network. Fibrous cartilage also has a dense network of fibers, but these fibers are not elastic. This
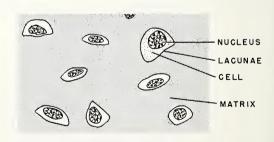
Fig. 10.5. Cartilage.

type of tissue is extremely strong and is utilized to connect ligaments and tendons to bone.

Calcification of cartilage normally takes place in bone formation. But it also occurs in old cartilage, and is then considered to be abnormal—to be a sign of cartilaginous degeneration. If cartilage is injured, regeneration is very slow, and in some cases it does not occur at all. In such cases the injured area simply becomes invaded by fibrous connective tissue.

Bone

Bone, known as **osseous tissue,** is one of the hardest of all the tissues. It has this quality because of the deposition of inorganic salts, primarily calcium phosphate, within the ground substance. Bone is living tissue; metabolic processes are constantly forming new bone and causing the dissolution of old bone.

If a long bone is cut longitudinally, it is found to contain a medullary canal filled with **bone marrow** (Fig. 10.6). In the center of the canal the marrow is generally yellow, while at the ends of the bone
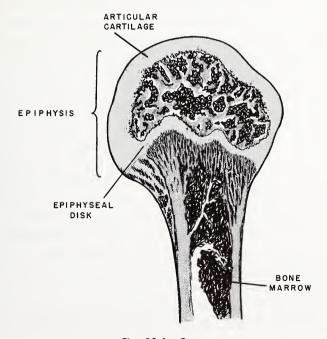
Fig. 10.6.   Bone.

it appears red. The red marrow is an important source of many of the blood cells. The bones, except at the ends where they are in articulation with other bones, are covered by a fibrous membrane called the **periosteum**. This membrane is thought to be essential for the repair of the underlying bone following injury. If bone is fractured, cells known as **osteoblasts**, which are located in the inner layer of the periosteum, cause new bone to be deposited between the fractured ends.

The formation of bone is a complex process. In the very young human embryo the skeleton is represented by cartilage and fibrous membranes. When the embryo is approximately eight weeks old, bone begins to develop in these two tissues. In other words, there are two sources of bone: 1) endochondral and 2) intramembranous.

Most of the bones of the body are formed from cartilage, that is, by endochondral ossification. The cartilage resembles a model of the future bone in that it has the same general shape and proportions. These cartilaginous precursors are covered by fibrous connective tissue called **perichondrium**. The shaft, or central part, is called the **diaphysis**. Each end is termed the **epiphysis** (Fig. 10.6). In the perichondrium there are fibroblasts which become osteoblasts. As the term indicates, the osteoblasts are responsible for the new formation of bone. When they become active, a ring of bone is deposited around the diaphysis. Once this ring is formed the perichondrium is properly called the periosteum, that is, it surrounds the bone. Blood vessels enter and spread throughout the cartilage which continues to grow while bone is being laid down. Eventually the cartilage, which is now surrounded by bone, disintegrates and bone takes its place, leaving only the central marrow. At this point there is a well-defined cartilaginous disk which separates the epiphysis from the diaphysis. Ultimately blood vessels enter the epiphysis, and bone is laid down extending from within out. Finally, at about the twentieth year of life, the cartilaginous disks which separate the epiphysis from the diaphysis disappear and bone development is complete. After this time the bones are incapable of increasing in length; thus growth of the individual is complete.

The bones of the face and cranium form by intramembranous ossification; there is no cartilaginous model. Instead, cells of the

membranes send out strands to form a framework. Calcium salts are then deposited in this framework to form bone, and the membrane now becomes the periosteum. When the bone is complete, it is found to consist of an outer layer of very compact bone surrounding an inner core of bone of a more spongy nature.

It has already been mentioned that there is a constant "turnover" of bone. The two types of cells primarily responsible for the formation and dissolution of bone are the osteoblasts and the **osteoclasts.** The osteoclasts are thought to function by causing the disintegration of the organic matrix of bone which then permits the deposited salts to reenter the body fluids.

Blood

Blood consists of the plasma and the cells. Since the cells are suspended in the fluid plasma, blood is very similar to other types of connective tissue which is characterized by the relative abundance of the ground substance.

Fresh blood appears brilliant red, thick, opaque, and homogeneous. It is about five times more viscid than water and has a specific gravity of about 1.055. Actually, blood is not homogeneous. If a thin film is placed under a microscope, its heterogeneous character becomes obvious. When blood is allowed to stand or is centrifuged, it separates into two distinct fractions: 1) the plasma and 2) the formed elements, consisting of red blood cells, white blood cells, and platelets.

Red Blood Cells. The red blood cells, also referred to as **erythrocytes,** are biconcave disks (see frontispiece, Plate 1). The red cell contains a nucleus only during its early formative period. In the blood stream, the cell is without a nucleus. In man each cubic millimeter of blood normally contains about 5,000,000 erythrocytes.

The red cells contain the complex protein-iron pigment, **hemoglobin.** Hemoglobin readily unites with oxygen and is thus essential for the transport of oxygen by the blood. Normal blood contains about 15 gm of hemoglobin per 100 ml.

Cells which give rise to erythrocytes compose what is known as **hematopoietic tissue;** red bone marrow is the most important hematopoietic tissue. The red cell passes through several stages before it reaches the mature form found in the circulation. Once in the blood

stream the red cell functions for a period of time and then disintegrates. The remains are removed from the blood by the liver and spleen. The life span of an erythrocyte in man is believed to average about 80 days. It has been calculated that 2 to 10 million red cells disintegrate per second. This means, of course, that a comparable number must be formed in the same period of time.

White Blood Cells. The white blood cells are also known as **leukocytes.** There are normally about 6,000 to 8,000 of these cells per cubic millimeter of human blood.

Various types of leukocytes are shown in Plate 1. They differ in the shape of the nucleus and in the character and staining qualities of the cytoplasm. **Neutrophils** make up about 60 to 70 percent of the white cells; the **lymphocytes** about 20 to 30 percent. Other leukocytes are the **monocyte, basophil** and **eosinophil.**

The bone marrow is the site of origin of the neutrophils, eosinophils, and basophils. These white cells and the erythrocytes may develop from the same parent cell. The white cells, unlike the erythrocytes, continue to mature and differentiate after they enter the blood stream. They retain their nuclei and are capable of further development. The lymphocytes and monocytes are formed in the lymphoid tissues, primarily in the lymph nodes and tonsils.

The best known role of the white cell is to aid in combating infectious processes. They function primarily by phagocytosis. The neutrophils are particularly valuable in this respect. In some manner the cells are attracted to the site of infection; such attraction is termed **tropism.** Once there, they engulf the infectuous agent and thus render it harmless.

Blood Platelets. The platelets of the blood are also termed **thrombocytes.** Many authorities do not consider them to be cells. Actually they are but fragments which possess no nucleus. Yet they are quite important. They contain substances which are essential for the formation of a blood clot, and they also become enmeshed in the developing clot so as to plug openings in the vascular system. The number of platelets normally varies between 300,000 and 600,000 per cubic millimeter of blood.

Plasma. The plasma is really the ground substance of the blood tissue. It occupies over half of the total blood volume. It is a straw-colored fluid composed of about 91 percent water and 9 percent solids.

The solids include the plasma proteins, **albumin, globulin** and **fibrin-ogen.** In addition there are the substances listed in Table 10.2.

TABLE 10.2. Partial Composition of Blood

Substance	Whole Blood	Plasma	Erythrocytes
		Grams/100 ml	
Water	78.0	91.0	65.0
Total solids	20.0	9.0	34.9
Organic	21.2	8.5	33.0
Inorganic	0.8	0.9	0.7
Total protein	18.5	7.0	30.0
Albumin		4.2	
Globulin		2.5	
Fibrinogen		0.3	
		Milligrams/100 ml	
Potassium	210.0	18.0	410.0
Sodium	190.0	320.0	35.0
Chloride	300.0	360.0	190.0
Calcium	5.5	10.0	
Iron	46.0	0.12	100.0
Total P	45.0	11.0	74.0
Inorganic P	5.0	3.8	6.0
Total N	3200.0	1000.0	5300.0
N.P.N.	35.0	27.0	47.0
Urea	24.0	25.0	22.6
Creatinine	1.3	1.5	0.8
Creatine	0.4	0.5	0.3
Total lipid		530.0	
Cholesterol	190.0	170.0	220.0
Lecithin	310.0	180.0	410.0
Glucose	70.0	80.0	65.0

MUSCLE

There are three types of muscle: 1) skeletal, 2) cardiac, and 3) smooth. These differ both histologically and physiologically.

Skeletal Muscle

Skeletal muscle, as the term suggests, is attached to the bones of the skeleton, making possible bodily movement. Because of the char-acteristic cross-striations, it is also known as **striated** muscle.

The basic unit of skeletal muscle is the **fiber,** sometimes referred to as a **myone.** The fiber is relatively long and contains one or more nuclei (Fig. 10.7). It is enveloped by a tubular sheath, the **sarcolemma,** which is an elastic, transparent, homogeneous membrane. Groups of fibers are held together in bundles by connective tissue. These bundles are termed **fasciculi.** A whole muscle consists of many fasciculi all bound together by connective tissue. In man, skeletal muscle fibers may be as long as 4 cm, but the diameter rarely exceeds 100 microns. The cross-striations are produced by alternating light

SKELETAL   SMOOTH   CARDIAC
MUSCLE   MUSCLE   MUSCLE
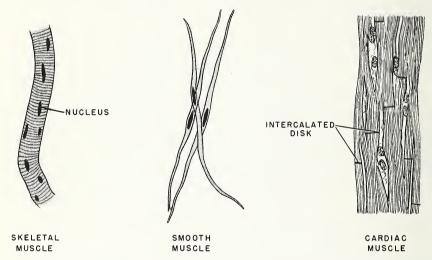
Fig. 10.7. The Three Types of Muscle. Individual skeletal and smooth muscle cells are shown. Note how in the cardiac muscle the cells form a syncytium.

and dark areas formed by longitudinally arranged **myofibrillae.** Thus, if a cross-section of a fiber is prepared, the myofibrillae will be cut. In such a preparation they are seen to be separated by the **sarcoplasm.** Each myofibril is composed of many actomyosin filaments. These filaments, as already explained (page 110), are the elements responsible for contraction of the muscle.

Cardiac Muscle

Whereas the skeletal-muscle fibers are distinct, clearly delineated by the sarcolemma, the fibers of cardiac muscle are seen to fuse into one another (Fig. 10.7). The limits are vague and the numerous nuclei seem to bear no relationship to individual fibers. That is to

say, cardiac muscle is a **syncytium,** a term that implies that the cells are together.

Cardiac muscle has striations but they are somewhat less distinct than in skeletal fiber; however, they are formed in the same manner. Cardiac muscle is characterized by the presence of **intercalated discs** which stain deeply and are thus quite prominent in fixed preparations. No function has yet been discovered for these discs.

Smooth Muscle

The smooth muscle fiber does not have cross-striations and each cell possesses only one nucleus. These fibers are fusiform or spindle-shaped. The nucleus is seen to lie in the central area (Fig. 10.7). The cells may be as long as 200 microns and as wide as 8 microns. There is one exception, namely the pregnant uterus in which the smooth muscle fibers become much larger.

Smooth muscle contracts very slowly, but the contraction can be maintained for long periods of time; in contradistinction, skeletal and cardiac muscle contract much more rapidly. Skeletal muscle can be activated at will and does not contract spontaneously. If it is denervated, it does not contract at all and therefore atrophies. Cardiac muscle and smooth muscle, on the other hand, are modified by their innervation, but even if completely denervated they continue to contract spontaneously and consequently do not atrophy.

Smooth muscle is located in the walls of the blood vessels, the respiratory passages, the alimentary canal, and the genito-urinary tract. It is also found in the skin, the eye, and in many glands.

NERVE

The **neuron** is the structural and functional unit of the nervous system. It consists of a cell body with one or more extensions, termed **processes.** Neurons are classified as multipolar, bipolar, or unipolar in accordance with the number of processes. A **nerve** is made up of many neurons held together by a connective-tissue sheath.

Multipolar Neurons

As the term indicates, the multipolar neuron has many processes. There is usually only one **axon,** but several **dendrites** (Fig. 10.8). In

the intact organism the impulse is usually propagated in only one direction and that direction is from the dendrites to the cell body and then along the axon.

DENDRITES

CELL BODY

NUCLEUS

CENTRAL PROCESS

AXON

AXON

CELL BODY

MYELIN

SCHWANN NUCLEUS

NEURILEMMA

NODE OF RANVIER

PERIPHERAL PROCESS

MULTIPOLAR

UNIPOLAR
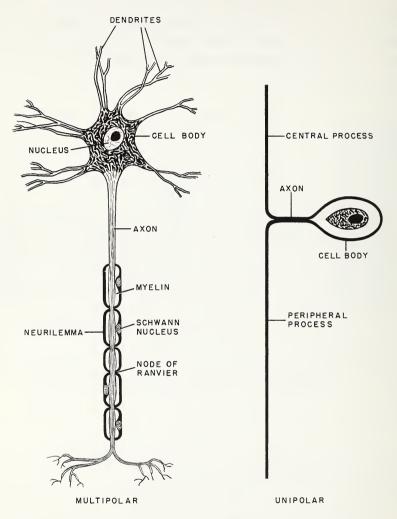
Fig. 10.8. Types of Neurons. Both processes in the unipolar neuron have the appearance of an axon. They are usually myelinated.

Bipolar Neurons

The bipolar neuron has but two processes which protrude from the cell body in opposite directions. The two processes are very simi-

lar structurally and bear more of a relationship to an axon than to dendrites.

Unipolar Neurons

Although the unipolar neuron has but one process coming off of the cell body, that process divides close to the cell body into two branches (Fig. 10.8). Again these processes have structural similarities to an axon. Thus, in a sense dendrites can be considered to occur only in the multipolar neuron. Some confusion arises, however, because the terms dendrite and axon are often used in a functional way. In this terminology, a dendrite is any process that propagates an impulse to the cell body, and an axon is the process that propagates the impulse away from the cell body. If these definitions are used, then almost all neurons have a dendrite and axon.

Dendrite Structure

The dendrites of a multipolar cell appear to be merely extensions of the cell body. The dendrites undergo considerable branching so as to form a network. That they have the same structure as the cell body is emphasized by the fact that in the intact organism the terminal surfaces of a preceding neuron are in contact with the cell body as well as the dendrites. And apparently the neuron may be activated at either point.

Axon Structure

The axon differs considerably from the cell body and from the dendrites. It is generally a long, unbranched structure which arborizes only at its end. The axon usually is covered with one or more sheaths. The axon and its sheaths make up the **nerve fiber.**

As seen in Fig. 10.8, there is a thin outer covering called the **neurilemma** or **sheath of Schwann.** Between the axon and this sheath there is another layer termed the **myelin** or **medullary sheath.** These coverings are periodically interrupted by the **nodes of Ranvier.**

The nodes of Ranvier influence propagation of the impulse. At these nodes ionic movement between the outside and inside of the axon can occur far more easily than in areas covered by the myelin sheath. Accordingly, it has been shown that propagation occurs from node to node; this is termed **saltatory propagation.** The term saltatory

is derived from the Latin saltatorius which means to dance, or to leap. Clearly, the term quite vividly describes the propagation of the impulse as it leaps from node to node over the myelin units.

The myelin is a semi-fluid substance composed of cholesterol, phospholipids, cerebrosides and other substances; it apparently serves as an insulating material. Recent electron miscroscope and X-ray evidence suggests that myelin has considerable structure. It appears to form something of a "bed roll" around the axon. Thus, in addition to serving as an insulator, it may also have a protective function.

Degeneration and Regeneration

If a nerve fiber is transected the part still connected to the cell body continues to survive, but the severed part undergoes degeneration. First the axon disintegrates and then the myelin sheath follows leaving only the empty neurilemma sheath. Although the nerve cell and the stump of the fiber survive, there are definite changes. The cell body swells, the nucleus is displaced, and the staining characteristics are altered. Thereafter the cell body returns to the normal state and the stump undergoes regeneration; it grows forward to enter the empty neurilemma. Ultimately the entire neurilemma becomes filled with the axon and new myelin, and regeneration is complete.

The Synapse

The junction between two or more neurons is termed the synapse. The neurons are in close association but there is no protoplasmic continuity at the synapse. The terminal arborizations of the axon of one neuron are in intimate contact with the dendrites or cell body of the following neuron.

It is now generally accepted that when the impulse reaches the end of the axon, **acetylcholine** is liberated, which then brings about depolarization of the secondary neuron so that an impulse is propagated along its length. At the synapse there is another compound called **cholinesterase** which quickly inactivates acetylcholine and thus prevents it from stimulating the secondary neuron more than once.

Because acetylcholine can only be liberated in sufficient quantities by the terminal arborizations of an axon, the impulse can be transmitted in only one direction across the synapse. It has already been pointed out that an impulse may be propagated in either direction

in a neuron, but in the intact nervous system, due to one-way trans-
mission at the synapse the impulses are normally propagated from
dendrites to cell body to axon.

SUMMARY

An organization of similar cells is termed a tissue. The three funda-
mental germ layers in the developing embryo are: 1) **ectoderm,**
2) **mesoderm,** and 3) **entoderm.** From them, epithelial, connective,
muscular, and nervous tissue develop.

Epithelium consists of one or more layers of cells with scarcely any
intercellular substance. The cells may be **squamous, columnar** or
cuboidal. If there is but a single layer, the tissue is **simple;** if two or
more, **stratified.** Epithelial tissue makes up the secretory elements of
glands. **Exocrine glands** have a duct; **endocrine glands** do not. A
simple gland has an unbranched duct; **compound** glands have a
branching duct system. The body **membran**es are formed by epi-
thelium. There are two types: 1) **serous** and 2) **mucous.** Epithelium
is capable of repair following injury.

Connective tissue serves as a binding material. It has a great abun-
dance of **intercellular** or **ground substance,** but a varying number of
cells and fibers. **Areolar** connective tissue is very loose; **fibrous** con-
nective tissue is more dense. In **adipose** tissue there is a great deposi-
tion of fat. The compact, slightly elastic ground substance causes
cartilage to be remarkably strong. Three types are recognized: 1) **hya-
line,** 2) **elastic,** and 3) **fibrous. Bone,** known as **osseous** tissue, is
the hardest of all tissues. This is due to the deposition of calcium
phosphate in the ground substance. There are two sources of bone:
1) **endochondral** and 2) **intramembranous. Osteoblasts** cause bone
formation; **osteoclasts** bone resorption. Blood is also a connective
tissue in which the **plasma** is the ground substance. The formed ele-
ments include, **erythrocytes, leucocytes,** and **platelets.**

There are three types of muscle: 1) **skeletal,** 2) **cardiac,** and 3)
smooth. Skeletal and cardiac muscle are striated; smooth muscle is
not. The fibers of cardiac muscle form a **syncytium.** Denervation
paralyzes skeletal muscle. Smooth and cardiac muscle remain active
following denervation and undergo spontaneous contractions.

The **neuron** is the structural and functional unit of the nervous

system. They are classified as **multipolar, bipolar,** or **unipolar.** When **dendrites** are present they usually branch extensively. There is generally but a single **axon,** covered by **myelin** and the **neurilemma.** The myelin is interrupted regularly by **nodes of Ranvier,** which causes saltatory propagation. Although a severed axon degenerates, if the neurilemma persists, the remaining stump will grow into it and regenerate. The junction between two neurons is termed the **synapse. Acetylcholine** is responsible for the transmission of the impulse at the synapse.

Problems

1. Outline the fundamental differences between epithelium and connective tissue.
2. Where in man would the following tissues be found:

    a. Transitional epithelium   d. Adipose tissue
    b. Serous membranes   e. Squamous epithelium
    c. Areolar connective tissue   f. Cartilage

3. How does the development of the cranial bones differ from those of the limbs?
4. How does smooth muscle differ histologically and physiologically from skeletal muscle?
5. In an isolated nerve the impulse is propagated in both directions. In the intact nervous system the impulse is propagated generally in only one direction. Explain.
6. Define:

    a. Synapse   d. Epiphysis
    b. Saltatory propagation   e. Hematopoietic tissue
    c. Periosteum   f. Nerve fiber

7. What is the function of the following types of cells: a) erythrocytes, b) macrophages, c) platelets.

Additional Reading

Bailey, F. R., "Textbook of Histology," 14th ed., Williams and Wilkins, Baltimore, Md., 1958.

Fiore, M. S., "An Atlas of Human Histology," Lea & Febiger, Philadelphia, Pa., 1957.

Maximow, A. A. and Bloom, W., "A Textbook of Histology," 7th ed., W. B. Saunders Co., Philadelphia, Pa., 1957.

Remington, J. W., ed., "Tissue Elasticity," American Physiological Society, Washington, D.C., 1957.

CHAPTER 11

THE NERVOUS SYSTEM

THE NERVOUS SYSTEM permits awareness of the environment, it regulates movement, and controls and coordinates the activity of various organ systems. These activities include circulation, respiration, and alimentation. For this reason a consideration of the physiology of the nervous system precedes the discussion of these other functions.

The nervous system is composed of two major components: 1) the central nervous system and 2) the peripheral nervous system. The former includes the brain and the spinal cord; whereas, the latter embraces the cranial nerves, the spinal nerves, and the autonomic nervous system.

PERCEPTION

Perception may be defined as knowledge through the senses of the existence and properties of matter and the external world. The senses permit the acquisition of knowledge and they also serve for protection, for pleasure, and for the integration of various functions.

Receptors

A receptor is specialized nervous tissue sensitive to a specific change in the environment. A change in the environment is termed a **stimulus.** The body is replete with receptors, located in the skin, in the special senses, such as the eyes and ears, and in deeper parts of the body. It has been shown that receptors react specifically to a particular type of stimulus. Thus, there are receptors for temperature, for pain, for light, and for many other stimuli.

When a receptor is stimulated appropriately, it responds by a burst of activity, by firing. It is this firing that activates the neuron with which the receptor is associated. As a result, an impulse is propagated by that neuron. The stronger the stimulus, the more rapidly the receptor fires and the greater is the number of impulses propagated per unit time.

Afferent Neuron

The impulse, initiated by the receptor, is propagated by a neuron to the central nervous system. Such a neuron is termed afferent. Once it has arrived in the central nervous system, the impulse may then take a variety of routes. For perception the impulse must be propagated to the cerebral cortex.

The Spinal Cord

Impulses arising from receptors throughout the body are propagated by afferent neurons in spinal nerves to the **spinal cord.** A cross-section of the spinal cord shows that it is divided into **white matter** and **gray matter** (Fig. 11.1). The white matter is composed of nerve tracts. The gray matter consists primarily of nerve cell bodies. These cell bodies give rise to axons which then leave the cord to innervate muscles. They are referred to as **motor** neurons. Other cell bodies, termed **internuncial** or **interconnecting** neurons, provide axons which course only a short distance and serve to connect two neurons within the spinal cord. Finally, some of the cell
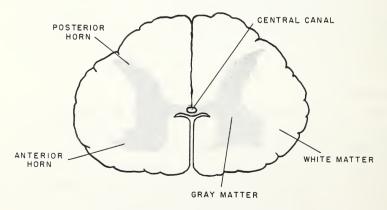
Fig. 11.1. Cross-section of the Spinal Cord.

bodies give rise to long axons which ascend the length of the spinal cord. These axons propagate impulses concerned with perception.

Major Sensory Pathways

There are two major sensory pathways. The **posterior columns** propagate impulses of touch, pressure, and proprioception. The **spinothalamic** tracts are concerned with pain and temperature sensation. The afferent neuron, propagating impulses of touch, pressure, or proprioception, enters the spinal cord and immediately turns upward to traverse the entire length of the spinal cord in the posterior columns without synapsing. In contradistinction, the afferent neurons propagating pain and temperature impulses end in the spinal cord. Secondary neurons then cross the cord to ascend on the other side.

The spinal cord is continuous with the **brain stem.** The lowest portion of the brain stem is termed the **medulla oblongata** (Fig. 11.2). In the medulla the neurons of the posterior columns end. Secondary fibers then cross the brain stem to ascend to the **thalamus.** The neurons that make up the spinothalamic tract go all the way through the brain stem to the thalamus. Here tertiary neurons arise which continue on to the **cerebral cortex.** In brief, both major sensory
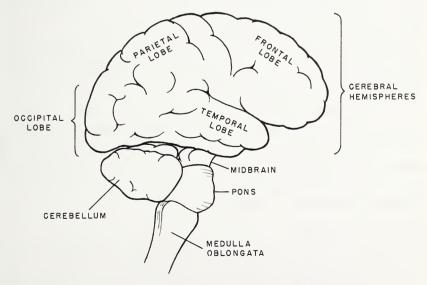
Fig. 11.2. The Major Subdivisions of the Brain.

pathways are composed of a chain of three neurons which cross to the opposite side of the central nervous system. It is for this reason that perception of a stimulus from one side of the body depends upon the cerebral cortex on the opposite side.

There is a similar pattern for impulses arising from receptors located in the head. The afferent neurons, in this instance, make up part of the cranial nerves. These neurons enter the brain stem to synapse with secondary neurons that cross and then ascend to the thalamus. Tertiary neurons propagate the impulses to the cerebral cortex.

The Sensory Cortex

The major portion of the brain is composed of the cerebral hemispheres; the outer covering of the hemispheres is termed the **cerebral cortex.** In man, the cerebral cortex reaches its greatest development. It is conveniently divided into four lobes: 1) frontal, 2) parietal, 3) occipital, and 4) temporal (Fig. 11.2). The tertiary neurons of the sensory pathways are projected from the thalamus to the cortex of the parietal lobe. For this reason it is called the sensory cortex. It has become the custom to designate areas of the cerebral cortex with numbers. The areas primarily concerned with perception in the sensory cortex are numbered 3, 1, and 2, usually designated 3-1-2. Areas 5 and 7 are secondarily associated with sensation.

If the entire sensory cortex is removed from both hemispheres, there is sensory deficit. In lower forms, there is considerable sensory retention which indicates that other areas of the brain are utilized for sensation in these species.

It is of importance to realize that the neurons have an orderly arrangement from receptor to sensory cortex so that one is able to localize with considerable accuracy the area of the body stimulated. Sensation can be localized on some parts of the body better than on others. For example, the finger tips are far more precise in this respect than is the back of the body.

Psychology of Perception

Perception connotes far more than the simple awareness of sensory impulses bombarding the sensory cortex; perception must be developed. The individual, through experience, comes to utilize mini-

mal sensory clues to identify objects. In other words, the simple awareness of a stimulus is sensation; the recognition of that stimulus is perception.

In many cases there are errors in perception. This is termed an **illusion.** However, **hallucination** is different. Whereas a false perception based on a stimulus is called an illusion, perception in the absence of a stimulus is an hallucination.

CONTROL OF MOVEMENT

It has been emphasized that skeletal muscle, which is necessary for bodily movement, depends upon its innervation. There are two basic types of movement: 1) voluntary and 2) reflex.

Voluntary Movement

The essential components of voluntary movement are: 1) a primary motor fiber which arises from a cell body in the motor cortex, 2) a secondary neuron which leaves the brain stem, or spinal cord, to innervate the muscle, and 3) the skeletal muscle.

Motor Cortex. Areas of the cerebral cortex concerned with voluntary movement are located in the frontal lobes (Fig. 11.2). These motor areas are designated as **area 4** and **area 6.** Area 4 is characterized by the presence of very large cells, termed **pyramidal** or **Betz** cells.

Motor Pathways. Axons from cells located in the motor cortex descend through the brain to end in association with the motor fibers of the various cranial nerves. These nerves innervate the muscles of the head. Other fibers from the motor cortex continue into the brain stem to end in synaptic union with short internuncial neurons which, in turn, synapse with the neurons that run in the spinal nerves to the muscles of the body. The first pathway is termed the **corticobulbar tract;** the second is the **corticospinal tract.** Whether or not there are also internuncial neurons interposed between the corticobulbar fibers and the cell bodies of the motor neurons of the cranial nerves is not known. The assumption is that there are. It is important to understand that most fibers that make up the corticobulbar and corticospinal tracts cross before terminating. Thus, the motor cortex on one side controls muscles on the opposite side.

Point-to-point Relationship. If area 4 is stimulated, it is found that specific muscles are caused to contract. These studies disclose that there is a point-to-point relationship between each part of area 4 and the various muscles. Thus, when the top of area 4 is stimulated, the muscles of the feet contract, and when the lower sections are stimulated, the facial muscles move. This means that there is an orderly arrangement from the pyramidal cells through the motor pathways to the muscles. Although some muscles are capable of more discrete movement than are others, this is not due to a difference in the muscles, but rather to the density of innervation, the so-called **innervation ratio.** Each motor neuron innervates a variable number of muscle fibers; the fewer the number innervated per neuron the more discrete is the movement. Thus very small areas of the lips may be moved independently, whereas only large regions of the buttocks can be activated.

Reflex Movement

A reflex is an involuntary response to a stimulus, a response that depends upon the integrity of the nervous system. The stimulus activates a receptor that fires the afferent neuron. The impulse, in the central nervous system is then transmitted to an efferent neuron, that is, one which propagates an impulse away from the spinal cord in this case, to a muscle which contracts (Fig. 11.3).

Pathways from Higher Centers. The reflex impulse is transmitted from the afferent to the efferent limb across one or more synapses, depending upon the complexity of the reflex. Synaptic transmission is influenced by the activity of higher centers. This influence is exerted by impulses which are propagated over pathways which descend the brain stem and spinal cord to end either on the efferent neuron, or on a small internuncial neuron which, in turn, ends on the efferent neuron of the reflex arc. Thus reflex activity may be enhanced, that is **facilitated,** or **inhibited.**

Spinal Shock. If the spinal cord is severed, all reflex responses below the level of the transection are lost for a variable period of time; a condition termed spinal shock. Ultimately spinal shock vanishes and reflex activity returns. Voluntary control, however, does not return. Thus, the muscle is paralyzed, but it will still contract reflexly.
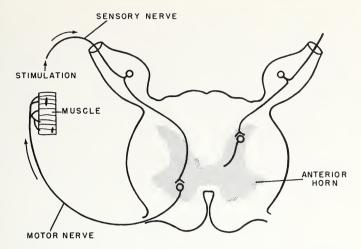
Fig. 11.3.   A Reflex Arc. The simplest reflex consists of a receptor which is stimulated, the sensory nerve, motor nerve, and the effector organ, in this case, a muscle. In many instances there is an inter-nuncial neuron between the sensory and motor neurons.

SPECIAL SENSES

The special senses include vision, audition, olfaction, and gustation. These sensory functions differ in many respects from the simple appreciation of pain and temperature, touch and pressure.

Vision

The eye functions in much the same manner as does a camera. In the camera the image is focused upon sensitive film by the lens. In the eye, there is also a lens which focuses the image upon the sensitive **retina** (Fig. 11.4). In the camera, there is a diaphragm to regulate the amount of light which enters the lens. In the eye, the **iris** has this function.

Focus of the Image. The lens of the eye is biconvex. Such a lens causes rays of light to converge to a focus. In a camera the rays are focused on the film by moving the lens backward or forward. In the eye, the lens is stationary but its shape changes so as to bend the rays a greater or lesser degree. The alteration in lens curvature in order to focus the image is termed **accommodation.** The curvature of the lens is under the control of the **ciliary muscles.**

DIAPHRAGM

LENS

FILM

IRIS

LENS

RETINA
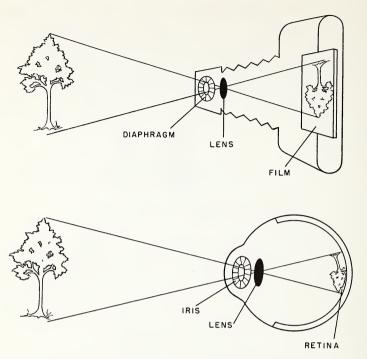
Fig. 11.4.   Comparison of a Camera and the Eye.

Visual Acuity. In order to see an object clearly, the object must be focused on the retina. This, as just indicated, is the function of the lens. In addition, the size of the object and the distance of the object from the eye play roles in visual acuity. In order for two objects to be differentiated as distinct entities, it is necessary for the resulting image of these objects to cover two or more light-sensitive receptors which are located in the retina. If the objects are so small or so far away that only a single receptor is stimulated, then objects will fuse together and appear as one. In practice, visual acuity is tested with standard charts known as **Snellen charts.** The standard distance from the chart to the eye is arbitrarily set at 20 feet. At this distance letters of a certain size can be clearly seen and identified with the normal eye.

Visual Receptors. The light-sensitive receptors in the retina are of two kinds: 1) **rods** and 2) **cones.** These receptors are histologically

distinct and they vary greatly in their sensitivity. The cones are not as sensitive as are the rods. The cones are utilized under conditions of relatively bright illumination; but in dim light, the rods are used. This arrangement permits the eye to have a tremendous range of light sensitivity. Relatively small changes in light intensity are quickly compensated by action of the iris. Greater changes involve a slower compensation which depends upon chemical reactions that take place in the rods.

The rods contain **rhodopsin,** sometimes referred to as **visual purple.** This substance responds to the radiant energy of light so as to activate the neuron associated with the rod. In bright light, rhodopsin bleaches and rod sensitivity is decreased. In low illumination, rhodopsin is reformed so that the rods become highly light sensitive. The synthesis of rhodopsin depends upon the presence of ample **vitamin A.** It is for this reason that individuals with vitamin A deficiency have night blindness.

Visual Pathways. The rods and cones initiate impulses in neurons which make up the **optic nerve.** These neurons end in synaptic union with secondary neurons in the lateral geniculate bodies of the brain. The secondary fibers then propagate the impulses to the cerebral cortex of the **occipital lobe.** Area 17 is the primary visual cortex, but areas 18 and 19 also are essential to normal visual function.

Color Vision. The normal eye can appreciate and differentiate various colors. Exactly how this is made possible is not completely clear. It is thought that the retina contains cones which are sensitive to specific colors. According to this hypothesis, the three basic colors of blue, red, and green stimulate specific receptors. The impulses are then propagated to the visual cortex, and here, depending on the number of receptors fired, mixing is done. As a result, the entire range of the color spectrum can be appreciated. Thus, if one looks at pure red, only one group of cones respond and by training the individual perceives red. The same is true of the other basic colors. But when a color that results from mixing two basic colors is viewed, then two types of receptors respond. The visual cortex, again by virtue of training, can interpret the resulting impulses as the mixed color.

Audition

The ear functions to gather and concentrate sound waves onto an area containing receptors sensitive to sound vibrations. The sound waves enter the external ear to impinge upon the **tympanic membrane.** The membrane, as a result, vibrates. As it moves back and forth, it causes three small bones of the middle ear to move. These are the **stapes, incus,** and **malleus** (Fig. 11.5). The stapes is attached to a small diaphragm that covers the **oval window.** Beyond the oval window is the inner ear which is filled with fluid. Thus, movements
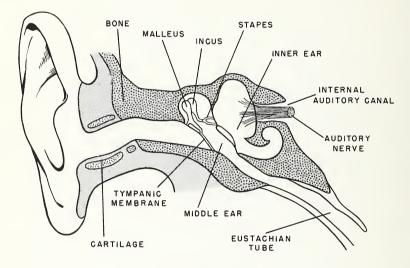
Fig. 11.5. Anatomy of the Ear.

of the oval window diaphragm set up waves in the fluid, and these waves then activate receptors which are located in the inner ear. In this way, sound waves ultimately activate receptors that initiate impulses in the neurons associated with the sound receptors of the inner ear.

Auditory Pathways. The neurons activated by the sound receptors make up the **auditory nerve.** They propagate the impulses to the medulla. From the medulla to the cerebral cortex the impulses are propagated over a series of neurons, some of which cross the brain stem, and others of which do not. The auditory cortex is located in the **temporal lobe** (Fig. 11.2).

Localization of Sound. A person with normal hearing can localize the direction from which sound waves emanate. The essential clue is the difference in arrival time of the sound at the two ears. Another clue used for localization is the difference in the intensity of the sound heard by the two ears.

Pitch Discrimination. The pitch of a sound is a function of the number of vibrations of the waves per second. The greater the number of vibrations the higher the pitch, or tone. In the inner ear, the sound receptors are located on the **basilar membrane** which roughly resembles a harp. It has been postulated that a sound of a particular pitch causes a specific area of the basilar membrane to respond more than any other area. By training, then, which permits one to localize this area according to pitch, one tone can be differentiated from another.

Intensity Discrimination. Exactly how one can differentiate a loud sound from a low one is not completely understood. A part of the mechanism seems to be the number of impulses per unit time that arrive at the auditory cortex. According to this viewpoint, a loud sound would activate more sound receptors than a weak one, and therefore more impulses per unit time would be propagated to the auditory cortex. However, since the number of impulses per unit time, that is, the frequency, is also thought to play a role in pitch discrimination, there is still more to be learned concerning these mechanisms.

Gustation

Gustation, or the sense of taste, is a chemical sense. There are chemical receptors located primarily on the tongue, termed **taste buds.** There appear to be at least four different types of taste buds, a fact which follows from the demonstration that man can easily distinguish substances that are: 1) sweet, 2) salty, 3) bitter, and 4) acid or sour.

Gustatory Pathways. Impulses initiated by the taste buds are propagated in three cranial nerves: 1) the **facial,** 2) the **glossopharyngeal,** and 3) the **vagus.** These neurons end in the medulla. Secondary neurons then cross the medulla to ascend to the thalamus. Tertiary fibers are projected to the sensory cortex of the parietal lobe.

Taste Discrimination. A wide variety of substances can be differentiated by the sense of taste. A part of this ability is due to the presence of the four types of taste buds which are sensitive to specific types of substances. The precise role that olfaction plays in taste discrimination is difficult to evaluate. Undoubtedly, the sense of smell does provide valuable clues in this respect.

Olfaction

Less is known concerning the physiology of olfaction than is true for the other senses. A wide variety of substances can be perceived by odor but the exact mechanism remains obscure. It is known that the odor receptors are in the olfactory epithelium which is located along the upper surface of the nasal passageways. These receptors activate neurons which make up the **olfactory nerves,** but central pathways and the parts of the brain essential for olfaction have not been adequately established.

THE AUTONOMIC NERVOUS SYSTEM

The autonomic nervous system is that part of the nervous system which innervates smooth muscle, cardiac muscle, and glands. It is, therefore, a motor system, but it is not under the control of the will. Accordingly, smooth muscle and cardiac muscle cannot usually be caused to contract voluntarily. Likewise, glands cannot be made to secrete in this way. Because these activities are involuntary and self-governing, the system is termed autonomic.

The autonomic nervous system differs from the somatic motor mechanism in that the autonomic fibers which originate in the spinal cord and brain stem do not innervate muscle or gland. These fibers end in synaptic union with a second neuron which does innervate the organ. Thus, there are **preganglionic** and **postganglionic fibers.** The preganglionic fiber terminates in a ganglion consisting of cell bodies which give rise to postganglionic fibers.

Sympathetic and Parasympathetic Divisions

The autonomic nervous system is composed of two parts: 1) sympathetic and 2) parasympathetic. The sympathetic fibers arise from the thoracolumbar segments of the spinal cord. The parasympathetic

fibers originate from the sacral segments of the spinal cord, and also in the brain to course with several cranial nerves.

Many organs of the body are innervated by both types of fibers. Stimulation of one division usually produces effects just opposite to those noted upon stimulation of the other member (Table 11.1). This provides very effective control over the function of the organ innervated.

TABLE 11.1. Autonomic Nervous System Control

Part	Sympathetic Action	Parasympathetic Action
Heart	Acceleration	Deceleration
Coronary arterioles	Dilatation	?
Eye pupil	Enlargement	Made smaller
Ciliary muscle	None	Contraction
Glands	Secretion	Secretion
Intestine	Decreased peristalsis	Increased peristalsis
Bladder	Uncertain	Contraction
Penis	Ejaculation	Erection
Cutaneous arterioles	Constriction	None

Visceral Reflexes

A visceral reflex is very similar to those already discussed; the only difference lies in the fact that viscera are involved instead of skeletal muscle. A visceral reflex consists of a receptor, an afferent neuron, one or more synapses, efferent neuron, and smooth muscle, cardiac muscle, or gland as effector organ. The efferent neuron in a visceral reflex is usually part of the autonomic nervous system.

Visceral reflexes control almost all the systems of the body. When circulation is discussed, it will be learned that heart action as well as blood pressure is controlled by visceral reflexes. The same is true of respiration and alimentation.

COORDINATION OF NERVOUS FUNCTION

Major aspects of the nervous system have been outlined. But it should be understood that these units do not function as independent entities; there is intricate coordination between all parts of the nervous system.

The Cerebellum

The cerebellum is an important center for coordination. It is a relatively large structure which lies just behind the larger cerebral hemispheres (Fig. 11.2). The cerebellum is connected by afferent and efferent pathways with many of the other parts of the nervous system. It is divided into three lobes by fissures: 1) the anterior lobe, 2) the posterior lobe, and 3) the flocculonodular lobe. The cerebellum may also be subdivided into: 1) the paleocerebellum, 2) the neocerebellum, and 3) the flocculonodular lobe. Paleocerebellum implies an old part, that is, one that was present very early in the development of the species. Neocerebellum, on the other hand, is a term to describe an area that is newer. The paleocerebellum is composed of the anterior lobe and a portion of the posterior lobe. The neocerebellum is comprised of the remainder of the posterior lobe.

The paleocerebellum seems to exert an inhibitory influence on lower motor neurons. This inhibition is simply an expression of coordination. During movement proprioceptors are activated and, as a result, impulses are rapidly propagated to the cerebellum. The cerebellum, in the light of this data and the demands of the cerebral cortex, can either enhance or inhibit the movement already in progress. The paleocerebellum seems to be the center for inhibition. In contradistinction, the neocerebellum has an excitatory, that is, a facilitatory function. Thus, by virtue of the neo- and paleocerebellum, movements may be facilitated or inhibited. But these influences are probably never initiated by the cerebellum. The cerebellum responds to the motor cortex and to the proprioceptors. In other words the cerebellum, due to its ability to facilitate or inhibit, can coordinate motor cortex drive in the light of the actual muscle movement. It also seems likely that the cerebellum not only inhibits or facilitates the lower motor neuron, but also fires directly to the motor cortex to change its excitability.

The flocculonodular lobe is essential to equilibration. There are sense organs associated with the inner ear which respond to changing positions of the body. These sense organs are the **semicircular canals** and the **utricle.** They initiate impulses which are propagated to the flocculonodular lobe for coordination with motor activity. In this

way, equilibration may be maintained under a wide variety of acts and conditions.

Basal Ganglia

The basal ganglia, also called subcortical nuclei, are extremely important centers of coordination. As the term subcortical indicates, these nuclei lie in the cerebral hemispheres below the cortex.

The clue to the function of the basal ganglia is provided by experiments in which the various nuclei are stimulated. If no movement is taking place, stimulation of the ganglia is usually without demonstrable effect. However, if movement is first elicited, for example by stimulating the motor cortex, then activation of the basal ganglia promptly inhibits that movement. From this and other lines of evidence it is generally believed that the major function of the basal ganglia is inhibition. Accordingly, the gross movement of muscles due to motor cortex activity can be selectively inhibited by the basal ganglia and thereby bring about orderly, smooth, and efficient muscular movement.

Associated Areas

The areas of the cerebral cortex necessary for motor and sensory function occupy a relatively small portion of the total cerebral cortex. The remainder constitutes the so-called association areas. These are regions of the cerebral cortex which are important centers of coordination.

Sensory Association Areas. The sensory areas are located in the parietal lobes. Contiguous to each region there is an association area. All of these association areas meet in what has been termed the **angular gyrus** to coordinate the various forms of sensation.

Prefrontal Areas. In front of area 6 and extending forward the entire remaining distance of the cerebral cortex is a region termed the prefrontal area. These areas have been designated by numbers 9 through 13. The word "personality" best expresses the role of prefrontal area activity. The prefrontal area seems to exert an inhibitory function over other areas of the brain. This causes the individual to pause before responding, to evaluate all of the factors and consequences, to plan his response, his course of action. In some indi-

viduals this inhibitory function is so exaggerated he cannot reach a decision at all. He becomes incapable of responding, he is abnormally depressed. **Prefrontal lobotomy,** an operation in which the prefrontal areas are undercut so as to destroy the afferent and efferent neurons, results in dramatic improvement.

Broca's Area. Speech is an act that requires incredible coordination of a vast amount of sensory and motor function. To coordinate all of the muscles that take part in articulate speech is the role of area 44, Broca's area, located in the frontal lobes. If this area is destroyed, the muscles used for speech are not paralyzed; they can still be volitionally contracted, but they cannot be contracted with sufficient coordination to produce speech.

Brain Waves. So-called brain waves represent electrical manifestation of the activity of the cerebral cortex. This electrical activity is of very low voltage. It may be sufficiently amplified to be recorded by an instrument termed the **electroencephalograph.** The record so obtained is called the **electroencephalogram,** or **EEG.** It has been found that the brain wave patterns vary during different activities. They also vary in certain abnormal brain conditions and thus are of value diagnostically.

Sleep. Sleep seems essential to the individual's well-being, and even to survival. Just why this is so is not clear and even less is known about the mechanism of sleep. It is now generally believed that the **hypothalamus** represents a wakefulness center. In other words, activity of the hypothalamus awakens the organism. In the absence of positive stimuli from the hypothalamus, the organism sleeps. The question still remains, however, as to what activates the hypothalamus and what accounts for the rhythm of sleep and wakefulness. It has been assumed that first the hypothalamus becomes active. This causes impulses to be propagated to various parts of the nervous system so that other activities occur. As a result of this activity, impulses are fired back to the hypothalamus. A self-regenerative cycle is set up so that full wakefulness occurs. To go to sleep, activity is reduced to a minimum which decreases the input to the hypothalamus. It becomes quiescent and decreases its output. Sleep occurs. It must be clearly understood that this is merely an oversimplified hypothesis. Sleep is undoubtedly far more complex than here presented.

SUMMARY

Receptors initiate impulses which are propagated by **afferent neurons** to the central nervous system. In the central nervous system the impulse may be propagated by sensory pathways to the cerebral cortex for perception, or it may be utilized to fire an **efferent neuron** which propagates the impulse back to a muscle or gland for reflex activity. The **posterior columns** propagate impulses of touch, pressure, and proprioception. The spinothalamic tracts are concerned with pain and temperature sensation. The regions of the cerebral cortex essential to general perception have been designated areas 3-1-2, and 5 and 7.

Movement is caused by skeletal muscle which depends upon its innervation. Movement may be voluntary or reflex. The motor cortex, areas 4 and 6, initiate voluntary movement. The **corticobulbar tract** and the **corticospinal tract** are the major motor pathways. Reflex activity may be either **facilitated** or **inhibited** by impulses from higher centers.

The **special senses** include vision, audition, olfaction, and gustation. There are specific receptors which initiate impulses propagated by cranial nerves to the cerebral cortex. For vision the receptors are the **rods** and **cones** of the retina. The occipital cortex is located in the occipital lobe. For audition, the receptors are located in the inner ear. The auditory cortex is located in the temporal lobe. The **taste buds** are the receptors for taste. Impulses are propagated to the sensory cortex in the parietal lobe. Olfactory receptors are located in the olfactory epithelium in the nasal passageways. Little is known of the central pathways or the essential area of the brain.

The **autonomic nervous system** is a motor system that innervates smooth muscle, cardiac muscle, and glands. It is divided into the **sympathetic** and **parasympathetic** divisions. In both there are **preganglionic** and **postganglionic fibers.**

The **cerebellum, basal ganglia,** and **association areas** function to coordinate nervous activity. **Broca's area** coordinates speech. **Brain waves** represent electrical manifestations of cerebral cortex activity. **Sleep** and **wakefulness** are associated with increased and decreased ac-

tivity of the nervous system as a whole. The **hypothalamus** seems to be a center for coordination in this respect.

Problems

1. An impulse is initiated by a pain receptor in the sole of the foot. Make a diagram showing the pathway over which this impulse will be propagated to reach the cerebral cortex.
2. Explain why destruction of the corticospinal tract results in paralysis yet the paralyzed muscles still respond reflexly.
3. On a sunny day you enter a motion picture theater. At first you can see nothing but the illuminated screen. After several minutes you can see the seats and the other people quite clearly. Explain the mechanisms responsible for this change.
4. What do the theories purporting to explain color vision and taste discrimination have in common?
5. How does the autonomic nervous system differ from the somatic nervous system?
6. What changes would you expect in an experimental animal in which: 1) the cerebellum has been removed and 2) the basal ganglia has been destroyed?

Additional Reading

American Physiological Society, "Handbook of Physiology—Section 1: Neurophysiology," 3 vols., Williams & Wilkins Co., Baltimore, Maryland, 1959-1960.

Buchanan, A. R., "Functional Neuro-Anatomy," Lea & Febiger, Philadelphia, Pa., 1957.

Eccles, J. C., "The Physiology of Nerve Cells," Johns Hopkins Press, Baltimore, Md., 1957.

Langley, L. L., Cheraskin, E., and Sleeper, R., "Dynamic Anatomy and Physiology," McGraw-Hill Book Co., New York, N. Y., 1958.

Netter, F. H., "Ciba Collection of Medical Illustrations," vol. 1, Ciba Pharmaceutical Products, Inc., Summit, N. J., 1953.

Ruch, T. C. and Fulton, J. F., "Medical Physiology and Biophysics," 18th ed., W. B. Saunders Co., Philadelphia, Pa., 1960.

THE CIRCULATORY SYSTEM

THE MAJOR FUNCTION of the circulatory system is to deliver oxygen and nutrient substances to all of the cells of the body. In addition, it transports metabolic waste products to the kidneys and to the lungs for removal. Also, hormones that are produced by the endocrine glands are carried by the blood to various sites where they exert their influence. The blood, as has already been pointed out, through the phagocytic action of some of its cells also serves for protection. Furthermore, in warm-blooded animals, the circulatory system is essential for the maintenance of body temperature.

To maintain pressure in the circulatory system and to keep the blood circulating, all of the three types of muscle are used. The heart, the main pump, is composed of cardiac muscle; in the blood vessels there is a layer of smooth muscle which can vary the caliber of the vessel and thus control pressure; and, as will be discussed, the veins are supported by the surrounding skeletal muscles. Contraction of these skeletal muscles serve importantly to propel blood through the veins back to the heart.

THE HEART

In man, the heart weighs about 350 gm. The principal constituent of the heart wall is the **myocardium,** made up chiefly of cardiac muscle. There are four cavities: two **atria** and two **ventricles.** There are four valves in the heart: 1) the **mitral valve** between the left

atrium and ventricle, 2) the **tricuspid valve** between the right atrium and ventricle, 3) the **aortic valve** between the left ventricle and the aorta, 4) the **pulmonary valve** between the right ventricle and the pulmonary artery.

Circulation of Blood Through the Heart

Blood enters the left atrium of the heart from the lungs; it then flows into the left ventricle which pumps it into the aorta for distribution throughout the body. The blood returns to the heart via the superior and inferior venae cavae which open into the right atrium. The blood then flows into the right ventricle to be pumped through the pulmonary arteries into the lungs. It returns via the pulmonary veins to enter the left atrium, thus completing the cycle (Fig. 12.1).
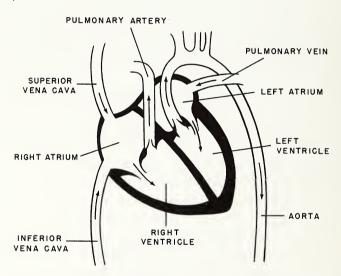
Fig. 12.1. The Heart Showing the Major Vessels and the Direction of Blood Flow.

The Heartbeat

Cardiac muscle, unlike skeletal muscle, has an inherent power of rhythmic contraction. It is true that the heart is liberally innervated by the autonomic nervous system which can modify the beat of the

heart, but there is a fundamental rhythm to the normal heart muscle which is completely independent of any innervation. The heart may be removed from an animal and perfused under optimal pressure with the proper constituents and apparently it will continue to beat indefinitely.

There are two phases of the heart beat: 1) the contraction phase termed **systole,** and 2) the relaxation phase called **diastole.** The heartbeat begins in the atria and then spreads so that in the normal beat the atria contract first and then the ventricles follow. In man the normal heart beats about 70 or 80 times per minute.

Pressure Relationships

The pressure of the blood as it enters the right atrium is very close to zero, that is, atmospheric pressure. The activity of the right side of the heart elevates the pressure to about 20 mm Hg. This pressure suffices to pump the blood through the pulmonary system. In the left side of the heart, the pressure is elevated to approximately 120 mm Hg, and this is the pressure at which the blood enters the aorta for circulation throughout the body.

It is the pressure of the blood that drives it through the blood vessels. Contraction of the myocardium, in combination with the action of the valves, is responsible for the build-up of pressure. During diastole the valves between the atria and ventricles are open, therefore, blood flows through the atria into the ventricles. Contraction of the atria just before the beginning of systole serves to pump more blood into the ventricle; then the powerful ventricles begin to contract. Since the outlet valves of the ventricles are closed, blood begins to flow back into the atria. But as soon as this happens the atrio-ventricular valves close, preventing this backflow. The ventricles are now closed cavities. The contraction of the ventricular muscle compresses the blood in these closed cavities and thus sharply elevates the pressure. Finally, the intraventricular pressure exceeds that in the pulmonary artery and in the aorta, so the outlet valves are forced open and blood is rapidly ejected under considerable pressure. After most of the blood is ejected, the ventricles cease to contract, intraventricular pressure falls, the valves close and the cycle begins over again.

The Heart Sounds

There are characteristic sounds associated with the beat of the heart that can be heard simply by placing one's ear close to the chest; however, for better localization, a **stethoscope** is used. For more detailed study of the sounds they may be converted into an electrical impulse by a microphone, amplified and then led to a loud-speaker, or to an ink-writing polygraph.

Usually, without the aid of amplification, only two sounds are heard. When the sounds are amplified, however, three, and often four, are detected. The **first sound** is heard at the beginning of systole and is caused mostly by the closing of the atrio-ventricular valves. The **second sound** comes at the end of systole and results from the closing of the aortic and pulmonary valves. The **third sound** is caused by the rushing of blood into the ventricle as soon as the atrioventricular valves open. And the **fourth sound** is associated with the contraction of the atrium.

Electrocardiography

In Chapter 7, page 139, it was mentioned that there is an action potential associated with the contraction of muscle. The heart is surrounded by an electrolyte medium capable of conducting the impulses of cardiac contraction to the surface of the body. If recording electrodes are attached to the surface of the body, the impulse can be led to suitable amplifiers for recording. The instrument used to amplify and record the electrical activity of the heart is called an **electrocardiograph;** the record is an **electrocardiogram.**

It is most convenient to attach the two electrodes to the wrists, or to one wrist and one ankle. A typical record is shown in Fig. 12.2. It will be immediately noted that there is a regular pattern. The pattern is repeated for each beat of the heart, thus the heart rate can be easily calculated. The electrocardiogram, abbreviated EKG, also quickly shows if the beat is regular or not.

The **P wave** is associated with the contraction of the atria, whereas the **QRS complex** is caused by ventricular activity. The **T wave** represents repolarization of the ventricles. By the analysis of the shape and amplitude of these waves and the intervals between them, considerable knowledge of normal as well as pathological activity of the heart is obtained.
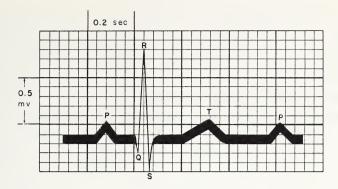
Fig. 12.2.   A Typical Electrocardiogram Showing the Principal Deflections. The P wave is associated with atrial contraction; the QRS complex with ventricular contraction (depolarization), and the T wave with ventricular repolarization.

Cardiac Control

The amount of blood ejected by each ventricle per minute is termed the **cardiac output;** whereas, the amount of blood ejected by one ventricle beat is called the **stroke volume.** The cardiac output, then, is equal to the stroke volume times the heart rate. The normal heart rate is about 70 beats per minute and the usual stroke volume is about 70 ml. Thus, in man, each ventricle pumps about 5 liters of blood per minute at rest; in severe exercise this may increase 6 or 7 times.

The output of the heart can be varied either by changing the rate of the beat, or the force of the contraction. It was pointed out in Chapter 6, page 121, that the greater the initial length of the cardiac fiber, the more forceful will be the contraction. The initial length of the cardiac fiber is determined by the quantity of blood that flows into the ventricle during diastole. If all other factors remain constant, then the force of contraction determines the quantity of blood ejected during systole. In short, stroke volume is directly related to end diastolic volume, up to a critical point (Fig. 12.3). This relationship which has been termed **Starling's law of the heart** is an inherent, self-regulatory mechanism that permits the heart to adjust to changing conditions.

Although the beat of the heart is an inherent property, the rate

and the force of contraction can be modified by the activity of the autonomic nerves that innervate it. The parasympathetic fibers reach the heart in the **vagus nerve;** the sympathetic fibers arise from the first five thoracic and the three cervical ganglia. Activity of the vagal fibers slows the heart and weakens the beat, whereas the sympathetic innervation has just the opposite influence. Thus, these two opposing forces can vary heart action over a broad range.
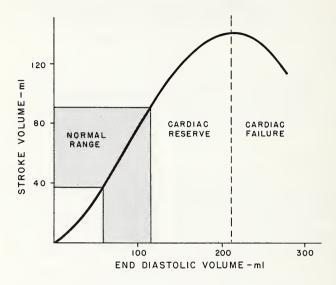
Fig. 12.3. Starling's Law of the Heart, Expressing the Relationship Between End Diastolic Volume and Stroke Volume.

Cardiac Centers. The dual control of the heart is coordinated by centers located in the medulla oblongata. These centers are composed of clusters of cell bodies. There are two centers, one that acts through the vagus to slow the rate and weaken the beat, and the other that functions through the sympathetic fibers to speed the heart and improve the force of contraction. Accordingly, they are termed the **cardioinhibitor center** and the **cardioaccelerator center.**

The centers serve to coordinate impulses that reach the medulla from various areas of the body. In this way heart function is keyed to body activity.

Reflex Control. At the bifurcation of each common carotid artery into the internal and external carotid arteries there is a swelling,

called a sinus, in the walls of which there are **pressoreceptors.** These are receptors sensitive to alterations in blood pressure. There are similar receptors in the arch of the aorta. Thus, one speaks of a **carotid** and an **aortic sinus.** An elevation in blood pressure causes the pressoreceptors to fire more frequently. The impulses are propagated to the medulla where they activate the cardioinhibitor center and depress the cardioaccelerator center. The result is a slowing of the heart and a weakening of the force of contraction. Cardiac output falls, and therefore so does the arterial blood pressure. The carotid and aortic sinus reflexes, in this way, regulate cardiac function. Through these reflexes an elevation of blood pressure results in decreased heart rate. This inverse relationship of blood pressure to heart rate is known as **Marey's law of the heart.**

In association with the carotid and aortic sinuses are clusters of specialized cells in what are called **bodies.** These cells, termed **chemoreceptors,** are sensitive to oxygen and carbon dioxide content of the blood. Activation of the chemoreceptors results in the propagation of impulses to the cardioaccelerator center which is activated, thus improving heart function. The chemoreceptors fire more frequently when the oxygen content of the blood is lowered. An elevation of carbon dioxide also has an effect, but a very weak one. The resulting improved cardiac output aids the situation by delivering more blood to the tissues and also by sending more blood through the lungs per unit time.

The reflexes initiated by the pressoreceptors and by the chemoreceptors are important self-regulating mechanisms. It is thought that there are other receptors in various parts of the circulatory system which function in a similar fashion to control cardiac function in accord with body demands.

ARTERIAL BLOOD PRESSURE

The circulatory system may be thought of as a closed circular tube. The heart is then the pump interposed between variable-sized arteries, veins, and capillaries that have different degrees of elasticity. The heart may eject a greater or lesser amount of fluid. This fluid, the blood, is a viscid substance. Thus, the circulation of blood involves the interrelationships between: 1) cardiac output, 2) caliber

of the vessels (cross-sectional area), 3) velocity of blood, 4) resistance to flow, 5) elasticity of the vessels, 6) blood viscosity, and 7) blood volume.

Circulatory Dynamics

At any point within a tube filled with a circulating fluid, force is exerted at right angles to the direction of flow. This force is termed **lateral pressure. Resistance** means opposition—opposition to the flow of blood.

Fluid does not flow through a tube as an intact cylinder; it flows in layers. The layers lying next to the vessel wall move very slowly, if at all. The next layer moves more rapidly and slides over the more slowly moving outermost layer. At the center of the tube, the flow is greatest. From these observations it can be understood that the smaller the diameter of a tube, the greater will be the resistance to flow. As a matter of fact, it has been shown that resistance to flow is inversely proportional to the **fourth** power of the tube diameter. In the circulatory system the diameter of the vessels can be altered by their smooth muscle layer. Small changes in the caliber of these vessels alters resistance, and therefore pressure, markedly.

The arterioles are primarily responsible for controlling the resistance to the flow of blood. Arterial blood pressure is measured upstream, so to speak, from the arterioles. To put it another way, the arterioles are downstream, or peripheral, to the arteries and thus the resistance they offer is spoken of as **peripheral resistance.** This is one of the two most important factors in controlling arterial blood pressure, the other being the cardiac output.

Circulatory System Volumes

It is important to understand that the volume of different parts of the circulatory system varies. The volume of any part is equal to the total cross-sectional area times the length. Although the capillaries have an extremely small cross-sectional area, there are so many millions of them that the total cross-sectional area is larger than any other part of the system. Nonetheless, they are so short that they hold only about 4 percent of the total blood volume. The volumes of the various parts of the circulatory system are given in Table 12.1. It should be noted that the venous system holds over twice as much blood as

do the arteries. Accordingly, variations in the capacity of the venous system can markedly alter circulatory dynamics, a fact that is too often overlooked.

TABLE 12.1. Capacity of Different Parts of the Circulatory System in Man

Part	Volume (ml)	Percent
Heart	250	5.0
Pulmonary system	1000	20.0
Arteries	750	15.0
Capillaries	200	4.0
Veins	2800	56.0
Total	5000	100.0

Systolic and Diastolic Pressures

During systole the aortic blood pressure rises to a peak and then falls (Fig. 12.4). The peak pressure is termed the systolic pressure. Throughout diastole there is a progressive fall in aortic pressure. Just before the aortic valve opens it reaches its lowest level. This
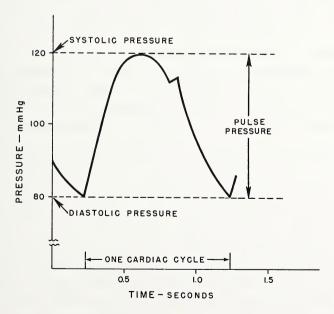
Fig. 12.4. Aortic Pressure.

level represents the diastolic pressure. The difference between the systolic and diastolic pressures is termed the **pulse pressure.** The average pressure during one cardiac cycle is called the **mean pressure.** An approximation of the mean pressure may be had by averaging systolic and diastolic pressures.

Influence of Elastic Vessels

The blood vessels are elastic tubes. The most elastic vessel is the aorta. If the aorta were distensible enough to enlarge proportionally with the volume of blood ejected by the heart during systole and without an increase in tension, there would be no increase in pressure. On the other hand, if the aorta were rigid, systolic pressure would rise very sharply and very high. The elasticity of the aorta, then, may be said to buffer the systolic pressure.

During diastole, the elastic aorta closes down upon the volume of blood as it flows into the larger arteries; therefore, the capacity of the aorta decreases with the decrease in volume. Were the aorta rigid the fall in pressure during diastole would be precipitous. The elasticity of the aorta, then, may be said to maintain the diastolic pressure. To put it another way, the magnitude of the pulse pressure varies inversely with the elasticity of the aorta and large vessels.

Determination of Arterial Blood Pressure

In both diagnostic and experimental work it is desirable to be able to determine the systolic and diastolic blood pressures. From these values the pulse pressure and mean pressure may be estimated.

The most direct method of determining blood pressure involves placing a cannula, or a large bore needle, directly into the artery. The cannula then is connected either to a mercury manometer or to a transducer for electronic recording.

Arterial blood pressure may also be determined by the use of a **sphygmomanometer.** This instrument consists of a cuff which is wrapped snugly about the arm just above the elbow. The cuff is connected to a manometer, and it is inflated until the pressure exceeds the arterial systolic blood pressure. The pressure is now slowly reduced while the examiner listens with a stethoscope to the sounds coming from the brachial artery distal to the cuff. When the pressure in the cuff falls just below systolic pressure, some blood will flow

under it and a sound will be heard. When this first sound is heard, the pressure is read from the manometer; this is the systolic pressure. Then, as the cuff pressure is allowed to decrease, suddenly the sounds change in intensity, or disappear altogether. This is the point of diastolic pressure.

Regulation of Arterial Blood Pressure

As has already been stated, the two major factors that regulate arterial blood pressures are the cardiac output and the peripheral resistance. The control of the heart was discussed above. The control of peripheral resistance will now be analyzed.

A decrease in the diameter of the arterioles is termed **vasoconstriction;** an increase is called **vasodilatation.** Vasoconstriction elevates arterial blood pressure; vasodilatation lowers it. The smooth muscle that regulates the diameter of the arterioles is under the control of the autonomic nervous system, but it is also influenced by various chemical substances.

Vasomotor Centers. Just as there are centers in the medulla to co-ordinate activity of the heart, so are there centers in the same area which coordinate **vasomotion.** Thus, one speaks of a vasoconstrictor and a vasodilator center. In recent years, however, the concept of a vasodilator center has lost favor. It is now believed that the medullary control of vasomotor activity is affected solely via variations in the vasoconstrictor discharge. It has therefore become the custom to speak of medullary **pressor** and **depressor** regions in order to make clear that they affect the vasoconstrictor tone and hence the arterial pressure by excitation and inhibition of the spinal vasoconstrictor neurons.

Reflex Control. The carotid and aortic sinuses initiate reflexes that regulate vasomotor activity as well as heart function. Accordingly, an elevation in arterial blood pressure evokes the following alterations: 1) slowing of the heartbeat, 2) decreased force of myocardial contraction, and 3) vasodilatation. The first two changes decrease cardiac output; the third decreases peripheral resistance. As a result, blood pressure falls. Decreased blood pressure evokes the reverse alterations. These reflexes, clearly, constitute important feed-back mechanisms to maintain blood pressure within a narrow range.

The carotid and aortic bodies respond to low oxygen tension of

the blood, elevated carbon dioxide and increased acidity. The end result of this response is to increase cardiac output and peripheral resistance. The blood pressure, therefore, increases.

Direct Chemical Control. Carbon dioxide acts directly on the arterioles causing dilatation. It also stimulates the vasoconstrictor center to produce vasoconstriction. These may seem to be opposing actions, but in fact they complement one another. Thus, when cells become active, carbon dioxide is produced, and the resulting vasodilatation permits more blood to flow to the active tissue. If the increased activity is great enough, then the CO_2 content of the blood may increase. If this occurs, the vasoconstrictor center will be activated. In the areas not directly dilated by high CO_2 there will be constriction and therefore blood will be shunted from these nonactive areas to the active ones. In this way large volumes of blood are forced to flow through the dilated vessels in the active tissue.

Epinephrine acts directly on the smooth muscle of most arterioles to cause vasoconstriction. Histamine has the opposite effect as does acetylcholine. In other words, epinephrine generally elevates arterial pressure whereas histamine and acetylcholine lower it.

VENOUS BLOOD PRESSURE

Although the circulatory system is a continuous series of vessels, it is nonetheless possible for very different conditions to prevail in various parts of the system. For example, in patients with chronic high blood pressure, the arterial systolic pressure can be well above 200 or even 300 mm Hg; yet the venous pressure may be within its normal range. Conversely, in congestive heart failure, the venous pressure may rise markedly with no change on the arterial side.

Determination of Venous Pressure

Venous pressures are so low that for greater accuracy they are measured by the use of water in place of mercury, or by very sensitive transducers. A rough estimate of venous pressure can be obtained simply by observing the neck veins. In the recumbent position these veins are not distended and usually are barely perceptible; however, if the right atrial pressure increases to over 10 mm Hg, then the neck veins are distended and become quite prominent.

Regulation of Venous Blood Pressure

The pressure in the veins, as in any system, depends upon the ratio of the input to the outflow. Blood enters the veins from the arterial system, and it flows out of the veins into the heart. Consequently, if the heart does not readily accept the blood from the veins, it will dam back and increase venous pressure. It is for this reason that high venous pressures occur in cases of a failing heart.

Venomotor Changes

As has been pointed out, the capacity of the veins is very large. If this capacity is decreased by constriction of the veins, termed **veno-constriction,** a large volume of blood will be forced toward the heart and the pressure will increase. Venomotor changes, by altering the volume of blood flowing to the heart, importantly regulate circulatory dynamics.

Skeletal Muscle

The veins are thin-walled vessels but they gain support from the skeletal muscle through which they course. When this skeletal muscle contracts, the vessels are squeezed. Since the larger veins have one-way valves, when they are squeezed, the blood can only move in the direction of the heart. Skeletal muscle activity, in this way, elevates venous pressure and increases the venous return to the heart.

Opposing the blood flow in the veins below the level of the heart, in the upright position, is gravity. If it were not for the support and pumping action of the skeletal muscles, gravity would cause the blood to pool in the veins. This sometimes occurs in individuals who stand motionless for long periods of time. Because too little blood is returned to the heart, the cardiac output falls, arterial blood pressure decreases, insufficient blood is pumped to the brain, and fainting results.

CAPILLARIES AND LYMPHATICS

A major purpose of the circulation is to deliver oxygen and nutrients to the tissues and to remove carbon dioxide and other products resulting from the metabolic processes. The interchange of these substances between the vascular system and the tissues occurs only

at the capillary level. The lymphatic system also serves to transport substances from the tissue spaces back into the circulation.

Capillary Wall

The capillary wall is extremely thin, and it is composed of but a single layer of endothelial cells. It is thought that in the capillary wall there are two sets of "pores": 1) small pores that allow passage of substances with a molecular weight not greater than 250,000, and 2) larger pores permitting much larger molecules to pass. Electron microscopy studies show the presence of intracellular vesicles packed tightly in layers in the capillary endothelium. It is thought possible that the large pores are these vesicles. The transfer of large molecules may consist of the material being taken up in a vesicle which then becomes pinched off. The term **cytopemphis** has been given to this process rather than **pinocytosis,** in order to convey the idea that substances are being transmitted through the cytoplasm rather than utilized by the cell, as is the case in pinocytosis.

Capillary-Tissue Space Interchange

The exchange of substances between the capillaries and tissue spaces is considered to be a purely passive process. That is to say, only physical forces determine the movement. Figure 12.5 shows the forces involved. It can be seen that the hydrostatic pressure tends to force fluid out of the capillary while the osmotic attraction of the nondiffusible proteins opposes this movement. Thus, in many capillaries there is movement of fluid out of the capillary at the arteriolar end and movement of fluid into the capillary at the end closest to the venule. If the hydrostatic pressure is very low, as it is in the capillaries of the pulmonary system, then there is movement of fluid into the capillary throughout its length. Conversely, if the protein content of the blood falls, as it does in malnutrition, then there may be excessive loss of fluid from the capillaries with the formation of **edema.** Edema, defined as excess fluid in the tissue spaces, also occurs in cases of venous or lymphatic obstruction.

The Lymphatics

The lymphatics are a network of fine vessels which begin as a series of closed tubes about the size of capillaries. These merge into
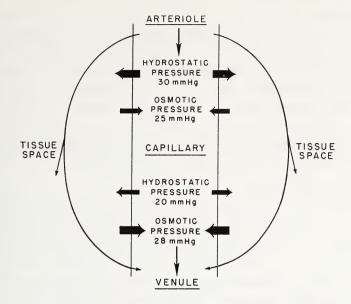
Fig. 12.5. Interrelationship of Pressures Responsible for Fluid Movement into and out of the Capillary.

larger and larger vessels. The lymphatics draining the lower extremities, lower trunk, and the left side of the chest form the **thoracic duct** which empties into the left internal jugular and subclavian veins. Lymphatics from the head, right arm, and right chest form the **right lymphatic duct** which empties into the junction of the right internal jugular and subclavian veins.

The main function of the lymphatics is to remove protein which makes its way to the tissue spaces.

Lymph Formation

By definition, only when fluid enters the lymphatic system does it become lymph. Accordingly, three fluids are involved: 1) blood plasma, 2) interstitial fluid, and 3) lymph. The interstitial fluid moves into the lymphatic capillaries to become lymph because the pressure in the tissue spaces is higher than it is in the lymph vessels. The lymph capillaries are so permeable that protein moves freely, thus there is no osmotic pressure to oppose this movement.

Lymph Flow

There is nothing comparable to the heart in the lymphatic system to cause lymph to move. It moves because the tissue pressure is higher than the pressure in the lymphatics, and it also moves because of the contraction of skeletal muscle and the pulsation of the arteries. There are one-way valves in the lymphatic vessels, therefore, when they are squeezed by skeletal muscle contraction, or by arterial pulsation, the lymph is caused to flow away from the tissue spaces toward the veins into which the lymphatic trunks empty.

Lymphatic System Function

One important function of the lymphatic system is to drain the tissue spaces. Excessive fluid formation is removed in this way, and in addition, the accumulation of protein in the tissue spaces is prevented. A second essential role is to combat infection. Foreign substances, such as bacteria, are engulfed by the polymorphonuclear leucocytes and macrophages. These cells are then removed from the lymph by the **lymph nodes.** In addition, the nodes seem to be responsible for the formation of **gamma-globulin** and specific **antibodies** to assist in combatting infection.

SUMMARY

The **heart** is the main pumping organ of the circulatory system. It is composed of cardiac muscle that has an inherent rhythm. This rhythm can be altered, however, by its innervation and by the influence of various substances on the myocardium. The two phases of the heartbeat are: 1) **systole** and 2) **diastole.** The blood enters the heart under very low pressure. It leaves the right ventricle at about 20 mm Hg, and the left ventricle at about 120 mm Hg. It is this pressure that drives the blood through the circulatory system.

The **electrocardiograph** is used to amplify and record the electrical activity of the heart. The record so obtained is called an **electrocardiogram.**

The amount of blood ejected by each ventricle per beat is termed the **stroke volume.** The amount of blood ejected per minute is called the **cardiac output.** In man, at rest, it is equal to about 5 liters. The heart is slowed and the force of contraction decreased by activity of

the **vagi nerves.** The sympathetic innervation has the opposite influence. The function of these nerves is coordinated by the **cardioinhibitor** and **cardioaccelerator centers** in the medulla. There are **pressoreceptors** and **chemoreceptors** in the **carotid** and **aortic sinuses** and **bodies** which monitor the blood pressure and composition. They initiate reflexes which control both cardiac function and arterial blood pressure.

The arterial blood pressure is regulated by: 1) cardiac output and 2) **peripheral resistance** exerted by the arterials. An increase in cardiac output or peripheral resistance elevates the pressure. Peripheral resistance is not only regulated by reflex mechanisms, but also by circulating epinephrine which, in most arterioles, causes vasoconstriction. Histamine and acetylcholine have the opposite effect.

Venous pressure is generally less than about 10 mm Hg. It is varied by: 1) the input from the arterial side, 2) the activity of the right side of the heart, and 3) venomotor changes. Gravity opposes the flow of blood in the veins below the heart in the upright position. These veins are supported by skeletal muscles which serve as auxiliary pumping mechanisms.

The interchange of substances between the vascular system and the tissue spaces occurs only at the **capillary level.** The capillary wall is extremely thin and permits diffusion of the plasma with the exception of the proteins. Hydrostatic and osmotic pressure determine the direction of movement of fluid into or out of the capillary. In most capillaries, fluid leaves at the arteriolar end and returns at the other end. Excessive fluid in the tissue spaces is termed **edema.**

The **lymphatic system** drains the tissue spaces, prevents the accumulation of protein in the tissue spaces, and serves to combat infection.

Problems

1. Draw a chart showing the pressure in the aorta, the left ventricle, and the left atrium during one complete cycle of the heart. Indicate where each valve opens and closes, and the point at which each of the four sounds can be detected.
2. A volume of fluid is suddenly introduced into the arterial system which elevates the arterial pressure. Discuss the sequence of events which would occur to lower the pressure to within the normal range.
3. A blood donor gives a liter of blood. Within a few hours his blood

volume is returned to normal due primarily to the movement of fluid from the tissue spaces into the capillaries. What are the mechanisms responsible for this movement?

4. In the aging process the elasticity of the aorta diminishes. Associated with this change the pulse pressures increases. Explain.
5. Vasoconstriction elevates arterial blood pressure; venoconstriction elevates venous blood pressure. Explain.
6. One of the cardinal signs of a failing heart is edema. Explain.

Additional Reading

Bernreiter, M., "Electrocardiography," Lippincott, Philadelphia, Pa., 1958.

Heymans, C., "Reflexogenic Areas of the Cardiovascular System," J. A. Churchill, London, England, 1958.

Langley, L. L. and Cheraskin, E., "Physiology of Man," 2nd ed., McGraw-Hill Book Co., New York, N. Y., 1958.

Luisada, A. A., ed., "Cardiology," McGraw-Hill Book Co., New York, N. Y., 1959.

Wiggers, C. J., "The Pressure Pulses in the Cardiovascular System," Longmans, Green and Co., New York, N. Y., 1958.

THE
RESPIRATORY
SYSTEM

RESPIRATION is sometimes divided into two categories: 1) **external** and 2) **internal** respiration. The latter is concerned only with the utilization of oxygen and the production of carbon dioxide and other metabolites by the cells. It is commonly referred to as cellular respiration and was discussed in Chapters 5 and 6. External respiration, on the other hand, includes the movement of air into the lungs, the passage of oxygen from the lungs into the pulmonary blood, the transportation of oxygen by the blood to all the cells of the body, the carriage of carbon dioxide by the blood to the lungs, the passage of this gas from the blood into the lungs, and the movement of air out of the lungs. In short, external respiration has to do with all the mechanisms that serve to deliver oxygen to the cells and to remove carbon dioxide from the cells and from the body.

MECHANISM OF BREATHING

The terms "respiration and breathing" are not synonymous. Strictly speaking, breathing is only a part of respiration; it is the part that has to do with the movements of air into and out of the lungs.

The Breathing Apparatus

In order to reach the lungs, air must pass through the nose or mouth, then into the pharynx, larynx, trachea, bronchi, bronchioles, and finally into the dead-end, sac-like structures, the alveoli. The alveoli have a very thin wall in intimate contact with the pulmonary capillary wall. It is across this double membrane that the transfer of gas occurs.

Although the mouth can be used as a respiratory passageway, the movement of the air through the nose is to be preferred. As a passageway, the nose warms, moistens, and because of the presence of cilia, filters the air. The movement of cilia in the nasal passages then moves the trapped particles into the pharynx. Although the mouth can moisten and warm the air too, it is not nearly as effective, and it has no means of filtration. In addition, mouth breathing leads to drying of the oral mucosa.

The Lungs. The lungs are extremely elastic and, when examined in the living state, feel somewhat like a sponge. They occupy the major part of the thorax. The outer surface of the lungs is covered by a fine, moist, glistening membrane, termed the **pleura.** It not only envelops the lungs but it also lines the inner surface of the thoracic cage. The area between the two pleural layers is referred to as the **intrapleural space.**

The right ventricle of the heart pumps blood into the pulmonary arteries. These arteries then arborize into progressively smaller branches until finally the capillaries of the lungs are reached. Although each capillary is minute, there are so many of them that their combined surface area is truly tremendous. This great surface area makes possible a rapid exchange of the respiratory gases.

The Thorax. The thorax, or **thoracic cage,** as it is often called because of the characteristic architecture of the ribs, is essentially a closed chamber. The ribs which are the framework for the thoracic cage are joined to each other by two sets of muscles: 1) the internal intercostal and 2) the external intercostal muscles. There are 12 pairs of ribs. The upper 10 pairs sweep around front and join, either directly or indirectly through cartilaginous connections to the **sternum.** The lower two pairs of ribs are free and are, accordingly, called floating ribs.

The **diaphragm** is a gently curving musculotendinous sheet which bounds the bottom of the thoracic cage. In its normal, relaxed position the diaphragm curves upward. However, when the diaphragmatic muscles contract, the diaphragm is straightened. In other words, the central portion which at rest curves upward, is now flat; this has the effect of enlarging the thoracic cage.

Thoracic Movements

There is a hingelike action of the ribs at their attachment to the spinal column. During inspiration the ribs swing upward. This enlarges the thoracic cage (Fig. 13.1). The muscles responsible for this
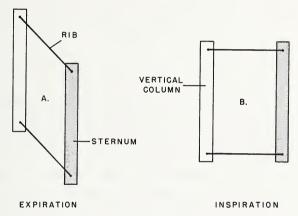
Fig. 13.1. Thoracic Cavity Model. During inspiration the sternum moves forward and upward. As shown in B, this enlarges the cavity.

movement, and therefore for inspiration, are: 1) the external intercostals, 2) the vertebral muscles, and 3) the sternocleidomastoids. The diaphragm, by descending also enlarges the thoracic cavity and contributes significantly to inspiration.

Expiration is usually a passive process. During this act, the muscles of inspiration relax allowing the anterior wall of the thoracic cage to move downward and thus its capacity is decreased. But expiration can be forceful. In this act there is not only relaxation of the muscles of inspiration, but also contraction of the muscles of expiration. These muscles of expiration include: 1) internal intercostals, 2) abdominal muscles, and 3) the pectoralis minor.

MOVEMENT OF AIR

During inspiration a volume of air moves through the passageways to enter the lungs. There some of its oxygen is removed and carbon dioxide from the body enters. During expiration this altered air leaves the lungs.

Inspiration

During inspiration the capacity of the thoracic cage is increased as just explained. At the same time the lungs are expanded for the following reasons:

1. *Surface Tension.* The two pleural layers are in intimate contact and they are moist. Thus, there is molecular attraction all along the surfaces of these layers which opposes separation of the layers. This surface tension is of considerable magnitude. Some idea of the magnitude of surface tension may be gained by attempting to separate two flat, wet plates of glass. The adhesive force of surface tension causes the lung wall to follow the thoracic wall.

2. *Intrapleural Pressure.* There is only a potential space between the two pleural layers. But even in the absence of surface tension the two layers would not separate very much during inspiration because as soon as the thoracic wall begins to pull away from the lung wall the pressure within the intrapleural space decreases. The pressure in the intrapleural space, the so-called intrapleural pressure, is normally below atmospheric pressure. Even at rest at the end of expiration the intrapleural pressure is about -3 mm Hg, that is, 3 mm Hg below atmospheric pressure. This is due to the fact that the thoracic cage is somewhat larger than the lungs, thus the lungs, even at the end of respiration, are stretched. The -3 mm Hg is a measure of their elasticity in the rest position. If an opening is made in the chest wall so as to permit air to enter the intrapleural space, the lung on that side will contract, i.e., collapse, and the intrapleural pressure will be equal to atmospheric pressure.

During quiet inspiration the expansion of the thoracic cage causes the intrapleural pressure to decrease from -3 mm Hg to about -7 mm Hg. The lungs are expanded because of the surface tension and the intrapleural pressure. The pressure within the lungs, i.e., the **intrapulmonary pressure,** decreases due to the enlarged volume. It is because the intrapulmonary pressure falls below atmospheric pressure

that air moves into the lungs. The difference between atmospheric and intrapulmonary pressures is the force that drives air into the lungs. The greater this difference the greater the volume and the more rapidly the air moves.

Expiration

During expiration the thoracic-cage capacity decreases. This permits the lungs, which are stretched during inspiration, to contract. As a result, both the intrapleural and the intrapulmonary pressures increase. The intrapulmonary pressure now rises above atmospheric pressure. It is for this reason that air moves out of the lungs. During passive expiration the intrapleural pressure never reaches atmospheric pressure, but during forced expiration the intrapleural pressure may rise as high as +40 mm Hg.

Respiratory Air Volumes

The average healthy individual has a maximal lung capacity of about 6 liters. This total is made up of the following components:

1. *Tidal Volume.* The quantity of air normally inspired with each breath while at rest is termed the tidal volume, or tidal air. It averages about 500 ml. As can be seen in Fig. 13.2 there is air in
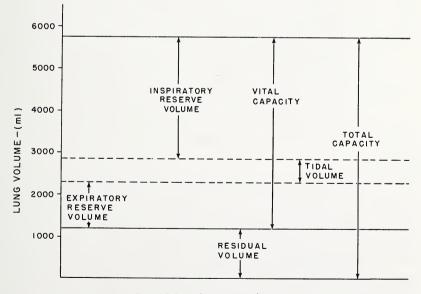
Fig. 13.2.   The Lung Volumes.

the lungs at the beginning of a normal inspiration, and room for additional air at the end of such an inspiration.

2. *Expiratory Reserve Volume.* By a forceful effort, at the end of a normal expiration, it is possible to expel a considerable volume of air, about 1,000 ml, from the lungs. This additional quantity is termed the expiratory reserve volume.

3. *Inspiratory Reserve Volume.* After a normal inspiration, one may continue to inspire and move an additional 3,000 ml of air into the lungs. The difference between the lung volume at the end of a normal inspiration and at the end of a maximal inspiration is termed the inspiratory reserve volume.

4. *Residual Volume.* Because the lungs cannot be completely collapsed by thoracic movements, there is still some air in the lungs even after a maximal respiratory effort. This quantity, termed the residual volume, amounts to about 1,200 ml.

Table 13.1 summarizes these respiratory air volumes. It can be seen that through a combination of maximal inspiratory and expiratory movement, approximately 4,500 ml of air can be expelled from the lungs. This volume, termed the **vital capacity,** includes: 1) tidal volume, 2) inspiratory reserve volume, and 3) expiratory reserve volume. It does not include the residual volume.

TABLE 13.1. Lung Volumes

Inspiratory reserve volume	3,000 ml
Tidal volume	500 ml
Expiratory reserve volume	1,000 ml
Vital capacity	4,500 ml
Residual volume	1,200 ml
Total capacity	5,700 ml

Ventilation

The quantity of air which passes through the lungs in one minute is termed the ventilation. It is the product of the rate of breathing and the volume of air drawn in during each inspiration. At rest, the respiratory rate is about 16 cycles (inspiration and expiration) per minute and the tidal volume is 500 ml. Therefore, ventilation at rest averages about 8,000 ml (16 × 500). The respiratory rate can be increased voluntarily and it is also regulated by other factors. It may

be increased as much as 8 to 10 times the resting rate, that is, up to well over 100 cycles per minute. By a combination of optimal rate and volume it is possible for the normal young adult to have a maximal ventilation of over 100 liters per minute.

TRANSPORT OF THE RESPIRATORY GASES

It was believed, at one time, that the epithelial membranes separating the alveolar air and the pulmonary blood possessed a secretory ability which served to move the gases; however, this concept is no longer held. The factors which determine the diffusion of the respiratory gases through these membranes are: 1) partial pressure of the gas, 2) permeability of the membranes, 3) chemical reactions in the blood, 4) rate of the pulmonary circulation, and 5) size of the alveolar surface area.

Transport of Oxygen

The partial pressure of oxygen in the lungs is about 101 mm Hg. The partial pressure of oxygen in the blood entering the lungs from the right side of the heart is about 40 mm Hg. It is this difference in pressure that causes oxygen to move from the lungs into the blood.

After oxygen reaches the blood, it must be transported to the tissue capillaries for transfer to the cells. The quantity of oxygen carried in physical solution is very small being on the order of about 0.2 ml per 100 ml of plasma. Blood, in contrast, can carry about 20 ml of oxygen per 100 ml of blood. It is obvious that the greater load of oxygen must be carried in chemical combination.

Hemoglobin. The characteristic color of blood is a function of an iron protein pigment termed hemoglobin which has a molecular weight of 66,700. Hemoglobin contains 4 Fe^{++} atoms per molecule. It is this ferrous iron that is responsible for the combination with oxygen. If the ferrous iron is oxidized to the ferric state, no oxygen is carried. The normal hemoglobin content in man is about 15 grams per 100 ml of blood. Each gram of hemoglobin is capable of combining with 1.34 ml of oxygen. Accordingly the **oxygen capacity** of the blood is about 20.1 ml (15×1.34), per 100 ml of blood.

The actual amount of oxygen in the blood is termed the **oxygen content.** The oxygen content of the arterial blood is normally close

to 20 volumes percent while the venous blood oxygen content may be only 15 volumes percent.

Oxygen Dissociation Curve. The ratio of the oxygen content to the oxygen capacity, expressed in percentage, is termed the **oxygen saturation.** The oxygen saturation of the blood is a function of the partial pressure of oxygen. This relationship is expressed by the oxygen dissociation curve (Fig. 13.3). It is important to observe that as the partial pressure of oxygen increases, more oxygen and hemoglobin unite until the hemoglobin is completely saturated with oxygen. It should also be noted that the degree of hemoglobin saturation with oxygen depends upon three factors: 1) the partial pressure of
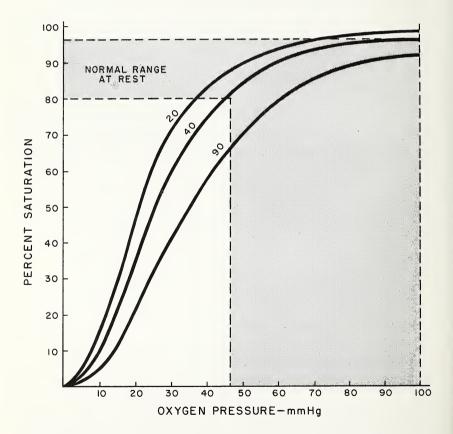
Fig. 13.3. Relationship of Oxygen Partial Pressure to the Percent Oxygen Saturation of the Blood Hemoglobin. The numbers on the curves indicate partial pressure of carbon dioxide.

oxygen, 2) the partial pressure of carbon dioxide, and 3) the pH, or acidity, of the blood.

The partial pressure of oxygen in the lungs is in excess of 100 mm Hg, therefore, as the blood flows through the pulmonary capillaries it becomes about 97 percent saturated. The partial pressure of oxygen in the tissues, at rest, is about 45 mm Hg. At this partial pressure the blood is only about 80 percent saturated. In other words, the blood gives up close to 20 percent of its load under normal resting metabolic conditions. An increase in metabolic rate lowers the oxygen partial pressure in the tissues and therefore more oxygen is dissociated from hemoglobin (Fig. 13.3).

Increased metabolic processes produce more carbon dioxide and acid metabolites. Both these conditions cause hemoglobin to give up oxygen. In summary, when the cells become more active, additional oxygen is quickly made available to them without alteration in circulatory or respiratory dynamics for three reasons: 1) lowered partial pressure of oxygen in the tissues, 2) increased partial pressure of carbon dioxide, and 3) increased acidity.

Transport of Carbon Dioxide

On the usual mixed diet, and at rest, the tissues produce carbon dioxide at a rate so that each 100 ml of blood leaves the capillaries bearing an additional 4 ml of CO_2. This volume of CO_2 is carried in the following ways:

1. *Combination with Hemoglobin.* Carbon dioxide combines with an NH_2 group in the hemoglobin molecule to form a carbamino acid, sometimes referred to as **carbhemoglobin** or **carbaminohemoglobin.**

2. *Formation of Bicarbonate.* About 75 percent of the carbon dioxide is carried as the bicarbonate ion. Before the CO_2 can enter the erythrocyte it must pass through the plasma where it reacts with water as follows:

$$CO_2 + H_2O \longrightarrow H_2CO_3 \longrightarrow HCO_3^- + H^+$$

In the plasma this reaction proceeds very slowly, therefore, much of the CO_2 continues on into the erythrocyte where the same reaction takes place, but at a very much greater rate. The reason for the dif-

ference in the speed of these reactions is the presence of a catalyst, **carbonic anhydrase,** in the erythrocytes.

3. *Physical Solution.* A small amount of carbon dioxide dissolves in the plasma and is carried in this way in physical solution.

Carbon Dioxide Dissociation Curve. In considering the transport of oxygen, the term oxygen saturation was used. Because oxygen is carried almost exclusively by hemoglobin, the degree of saturation of the hemoglobin with oxygen is an accurate measure of the quantity of O_2 carried. However, the CO_2 combined with hemoglobin represents only a part of the total CO_2 transported, therefore, the degree of hemoglobin saturation would not necessarily present an accurate estimate of the quantity of CO_2 carried. For this reason, the term carbon dioxide **content** is used which is the amount of the gas present, expressed in volumes percent.

It is seen in Fig. 13.4 that the CO_2 content varies with the partial pressure of the gas. At rest, the partial pressure of CO_2 in the arterial blood is about 40 mm Hg. From the dissociation curve it can be seen that the arterial blood contains about 50 volumes percent. In the venous blood the CO_2 partial pressure is normally 46 mm Hg and the content would be about 54 volumes percent.
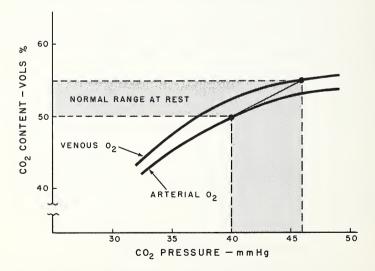
Fig. 13.4. Relationship of Carbon Dioxide Partial Pressure to the Carbon Dioxide Content of Blood. It can be seen that this content is also dependent upon the oxygen partial pressure.

Just as carbon dioxide influences the combination of hemoglobin and oxygen, so does oxygen affect the combination of carbon dioxide and hemoglobin. Accordingly, at least two local mechanisms aid in the removal of carbon dioxide from active tissues without alteration in cardiovascular or respiratory dynamics: 1) increased CO_2 partial pressure, and 2) decreased O_2 partial pressure.

Diffusion of Carbon Dioxide into the Lungs. The partial pressure of CO_2 in the blood entering the lungs from the right ventricle is about 46 mm Hg. In the alveolar air it is only 40 mm Hg. The differential between these two pressures is the force that drives the carbon dioxide from the blood through the membranes and into the alveolar air, to be expelled during expiration.

THE CONTROL OF BREATHING

At rest, about 8 liters of air move through the lungs per minute. This 8 liters contains about 1.6 liters of oxygen of which approximately 300 ml enter the blood. At rest, some 5 liters of blood flow through the pulmonary and tissue capillaries per minute. The arterial blood entering the tissue capillaries contains close to 20 volumes percent of oxygen. About 4 volumes percent are given up to the tissues. This suffices for the basal metabolic processes. But when the organism becomes more active, more oxygen is required. This need, if not too great, can be satisfied at the local level by a greater utilization of the available oxygen. If this fails to satisfy the needs, local vasodilatation results in greater blood flow and thus the deliverance of more oxygen per unit time. The next step, for even greater needs, is a general increase in blood flow, that is, increased cardiac output. And finally, the ventilation may be increased from the basal 8 liters per minute to over 100 liters per minute.

The Respiratory Center

The thoracic and diaphragmatic movements require a high degree of coordination. This coordination is under the control of a cluster of nerve cells in the brain stem which together constitute the respiratory center.

The respiratory center is located in the upper two-thirds of the medulla and extends on up into the pons. The cells which compose

this center end in synaptic union with cranial nerves and spinal nerves. By virtue of this arrangement the muscles of the face, throat, chest, and diaphragm are coordinated for respiratory purposes.

Chemical Control of Breathing

The cells that make up the respiratory center may be thought of as chemoreceptors, since they are responsive to the chemical state of the blood. These cells are more sensitive, by far, to carbon dioxide than to any other substance.

Carbon Dioxide. The influence of carbon dioxide on ventilation is shown in Fig. 13.5. There is normally only about 0.04 percent CO_2 in the inspired air. No measurable alteration in ventilation occurs until the inspired air contains at least 1 percent CO_2. When it contains 4 percent, ventilation is doubled, and at higher concentrations the ventilation increases very sharply until a maximum of about 80 or 90 liters per minute is reached. Most individuals can tolerate about 10 percent CO_2 in the inspired air, but higher concentrations produce great discomfort and then depression and uncon-
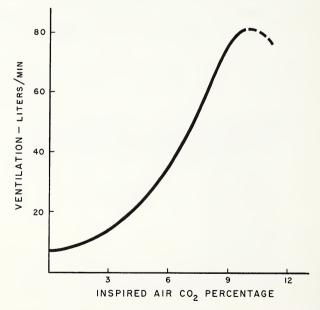
Fig. 13.5. Influence of Carbon Dioxide on Ventilation. Carbon dioxide in the inspired air in excess of about 10 percent becomes intolerable.

sciousness. Ultimately respiratory paralysis will result and therefore the curve, at that point, falls sharply.

Acidity. It is true that increasing acidity of the blood increases ventilation. But the quantitative aspects must be kept in mind; even in cases of severe acidosis the ventilation is only increased about 4 times.

Oxygen. A decreased oxygen content of the blood, that is **hypoxemia,** actually depresses the respiratory center. Thus, it can be concluded that carbon dioxide is the only true physiological stimulant of the respiratory center. In pathological conditions, acidosis also increases ventilation.

Reflex Control of Breathing

The respiratory center receives impulses from many areas of the body, including the cerebral cortex. Those from the cerebral cortex result in voluntary control of breathing. All others enter into reflex patterns.

Hering-Breuer Reflex. The cells that make up the inspiratory center either fire spontaneously, or in response to the carbon dioxide in the surrounding medium. Impulses are propagated to the inspiratory muscles, the thorax expands and so do the lungs. In the lungs, there are receptors sensitive to stretch. Accordingly, as the lungs expand, these receptors fire more and more rapidly. The impulses so initiated are propagated by afferent neurons in the vagal nerves to the respiratory center which, as a result, is inhibited. Thus, inspiration ends and expiration occurs. This reflex was described by Hering and Breuer in 1868 and still bears their name.

Pneumotaxic Reflex. In the pons there is a cluster of cells forming a center that has been termed the pneumotaxic center. The role of this center in respiration has been seriously questioned, but it is generally held that under some circumstances, the pneumotaxic center does function. For example, it is found that after the vagi are cut, breathing first becomes slower and deeper, as would be expected. But after a period of time, breathing returns to normal. Some investigators believe that the pneumotaxic reflex takes the place of the Hering-Breuer reflex under these and other conditions. Thus, it is postulated that when the respiratory center fires, impulses not only go to the motor neurons, but also to the pneumotaxic center. In

turn this center then fires impulses back to the respiratory center resulting in inhibition of inspiration and perhaps activation of expiration.

Carotid and Aortic Sinus Reflexes. As outlined in the previous chapter, there are pressoreceptors in the aortic and carotid sinuses. A rise in blood pressure activates these receptors and, as a result, there is slowing of the heart and a decrease in arterial blood pressure due to decreased peripheral resistance and diminished cardiac output. At the same time, there is also depression of breathing. Conversely, a fall in blood pressure causes hyperventilation. It is thought that at least a part of the hyperventilation which characterizes severe hemorrhage, and circulatory shock, is due to this pressoreceptor response.

Carotid and Aortic Body Reflexes. Stimulation of the nerves that innervate the carotid and aortic bodies results in effects opposite to that of activation of the pressoreceptors, that is, there is an elevation in blood pressure, in heart rate, and there is increased ventilation. From a quantitative standpoint it should be understood that the carotid and aortic sinus mechanisms are far more important in the control of circulation than in the regulation of breathing. On the other hand, the carotid and aortic body reflexes are of much greater significance to respiration than to circulation.

The chemoreceptors of the carotid and aortic bodies are activated by: 1) decreased oxygen, 2) increased carbon dioxide, 3) decreased pH, and 4) increased temperature.

In brief, it can be said that the chemoreceptors constitute the only mechanism for response to hypoxemia. In addition, they constitute a second line of defense to furnish respiratory drive when the primary mechanisms fail. When the chemoreceptors in the aortic and carotid bodies are denervated, ventilation decreases by about 30 percent. This would indicate that the chemoreceptors exert a tonic facilitating influence on the respiratory center.

Joint Reflexes. There are proprioceptors in the muscles and tendons which are activated by the movement of the joint. It has been shown that a very profound increase in ventilation results due to the propagation of impulses to the respiratory center from these receptors. Undoubtedly, in exercise this is a major reflex mechanism responsible for the characteristic hyperventilation.

Voluntary Control of Breathing

Breathing is unique in that it is controlled not only by exquisitely attuned automatic mechanisms, but it also is under voluntary control. However, the voluntary control is not absolute. That is to say, there are limits so that the breath cannot be held indefinitely, nor may hyperventilation be voluntarily continued for long periods of time.

The demonstration that breathing can be modified at will indicates that the respiratory center is influenced by the cerebral cortex. Stimulation experiments support this view.

The normal rhythm of breathing may be altered, or completely stopped voluntarily. A complete cessation of breathing is termed **apnea.** During the period of apnea carbon dioxide accumulates in the blood and becomes an ever greater stimulant to the respiratory center. Ultimately, this stimulant is great enough to overcome the volitional inhibition, and breathing once again occurs.

Control of Breathing during Exercise

One of the most fascinating, and yet perplexing fields of study is the response of the breathing mechanism to exercise. It is quite easy to demonstrate that ventilation is remarkably attuned to tissue demands, but the mechanisms responsible for this relationship are not completely understood. Figure 13.6 shows some of the more
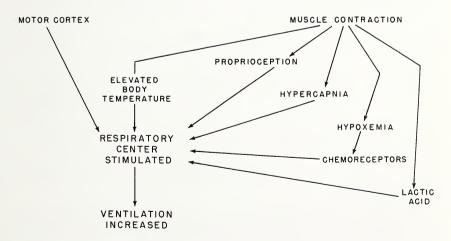
Fig. 13.6. Mechanisms Responsible for Increased Ventilation During Exercise.

important mechanisms that play roles in the increase in ventilation during exercise. It must be understood that only in exhaustive exercise does the blood concentration alter sufficiently to become a factor in respiratory drive. The major factors appear to be impulses from the motor cortex, and proprioceptive impulses from the actively contracting muscles.

SUMMARY

In the intact, multicellular organism, respiration involves the movement of air into the lungs, the diffusion of oxygen into the blood, the transport to the cells, the movement of **oxygen** into the cell, and the utilization of that gas by the cell. It also involves the production of **carbon dioxide** and other metabolites which must then be expelled from the body.

Air moves into and out of the lungs because of differentials between the atmospheric pressure and **intrapulmonary pressure**. The intrapulmonary pressure is varied by movements of the thoracic cage and the diaphragm. In man the total lung capacity is about 5,700 ml. In quiet breathing about 500 ml of air are moved per breath. With a respiratory rate of 16 cycles per minute, ventilation is equal to 8 liters per minute.

The partial pressure of oxygen in the lungs is higher than in the blood, therefore oxygen diffuses into the blood. It combines with **hemoglobin** avidly. The **oxygen capacity** of the blood is about 20 volumes percent. At rest, about 20 percent of this load is given up to the cells. The amount of oxygen dissociated from hemoglobin in the capillaries depends upon: 1) oxygen partial pressure, 2) carbon dioxide partial pressure, and 3) the pH.

At rest, on the usual mixed diet, the cells produce carbon dioxide at a rate so that each 100 ml of blood leaves the capillaries bearing an additional 4 ml of CO_2. In the blood the CO_2 is carried as bicarbonate, in combination with hemoglobin, and in physical solution. In the pulmonary capillaries the partial pressure of CO_2 is higher than it is in the alveolar air, therefore this gas diffuses into the lungs to be expelled.

Ventilation varies in accord with cellular needs for oxygen and the elimination of carbon dioxide. Breathing is under the control of

the **respiratory center** located in the medulla and extending as high as the pons. The cells that compose the respiratory center are chemoreceptors most sensitive to CO_2. They also are excited by lowered pH, but are depressed by **hypoxemia.**

The respiratory center can be modified by impulses from the cerebral cortex as well as from other areas of the body. Stretch of the lungs during inspiration evokes the **Hering-Breuer reflex** to terminate inspiration and start expiration. The **pneumotaxic reflex** has a similar function. Alterations in the arterial blood pressure and composition of the blood influence the carotid and aortic sinus and body receptors. As a result there are not only modifications in the circulatory system but also in respiration. A rise in pressure depresses breathing; however, hypoxemia increases ventilation, as does increased CO_2, decreased pH, and increased temperature. There are receptors in the muscles and tendons which can influence breathing. These play a major role in coordinating ventilation with exercise.

Problems

1. No matter how forceful the expiratory effort, it is impossible to expel the residual volume from the lungs. Why?
2. At 18,000 feet the atmospheric pressure is just one-half that at sea level. At 18,000 feet a sample of alveolar air is found to have the following composition: oxygen 12%, CO_2 11%, N_2 64%, water vapor 13%. Calculate the oxygen saturation of the arterial blood under these conditions.
3. If during exercise the carbon dioxide partial pressure in the tissues increases to 55 mm Hg, what would be the carbon dioxide content of the venous blood?
4. During severe exercise ventilation may exceed 80 liters per minute. Outline the mechanisms responsible for this increase in ventilation.
5. Define:

    a. Intrapleural pressure   d. Oxygen content
    b. Tidal volume   e. Hypoxemia
    c. Vital capacity   f. Volumes percent

6. In what units are the following usually expressed:

    a. Oxygen saturation   d. Carbon dioxide content
    b. Oxygen content   e. Ventilation
    c. Oxygen capacity   f. Respiration

Additional Reading

Campbell, C. J. M., "The Respiratory Muscles and the Mechanics of Breathing," Year Book Publishers, Chicago, Illinois, 1958.

Comroe, J. H., "The Lung," Year Book Publishers, Chicago, Illinois, 1955.

Fenn, W. O., "The Mechanism of Breathing," *Scient. Am.*, **202**, 138-148 (1960).

Guyton, A. C., "Textbook of Medical Physiology," 2nd ed., W. B. Saunders Co., Philadelphia, Penn., 1961.

THE URINARY SYSTEM

ONE OF THE functions of the respiratory system is elimination, the elimination of a major product of metabolism, carbon dioxide. Other substances are excreted from the body by the urinary system which consists of the kidneys, ureters, bladder and urethra.

The primary function of the kidneys is to regulate the volume and composition of the body fluids. This regulation is necessary because changes in the environment, in the food and fluids ingested, and in metabolism tend to alter the constancy of the internal environment, a constancy that has been termed **homeostasis**. The kidneys, by their ability to alter the volume and composition of the urine over a truly remarkable range, are able, in most instances to prevent serious alterations in the internal environment. In short, the kidneys are vital homeostatic organs.

ANATOMY OF THE KIDNEY

The kidneys are located high in the posterior part of the abdominal cavity. They are bean-shaped organs that weigh, on the average in man, about 170 grams each.

The Nephron

The functional unit of the kidney is the nephron. It has been estimated that there are about one million of these units in each kidney. Each nephron consists of a **glomerulus, a proximal tubule,** a

thin loop of Henle, and a distal tubule which then joins a collecting duct (Fig. 14.1). The glomerulus is formed by a capillary bed that is encased in Bowman's capsule.
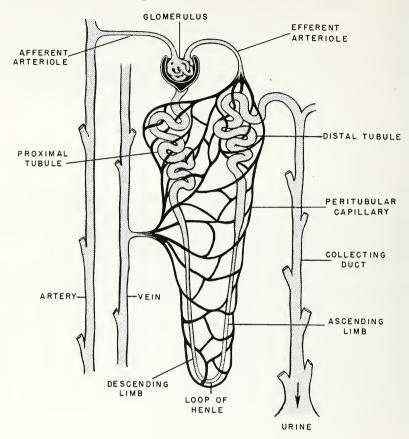
Fig. 14.1.   Anatomy of the Nephron.

Blood Supply

The circulation through the kidney is unique. Blood enters via the renal artery which springs directly from the aorta. This artery subdivides into progressively smaller branches. Ultimately, individual branches, called afferent arterioles, pass to each glomerulus to subdivide into a typical capillary bed. The capillaries then join to form, not a venule, but rather another arteriole, the efferent arteriole. This is the only example in the organism of a capillary network lying

between two arterioles. The efferent arteriole then subdivides into a capillary bed that surrounds and supplies the remainder of the nephron. These capillaries empty into a venule which joins others to form the renal vein. The blood is then returned to the inferior vena cava.

Both the afferent and efferent arterioles are well supplied with smooth muscle, and are thus capable of vasomotor changes. Because they can alter resistance independently of one another, the blood pressure within the glomerular capillary can be markedly changed.

THE FORMATION OF URINE

The first step in the formation of urine is the filtration of a part of the plasma through the capillary of the glomerulus to enter the tubule. The filtrate then passes through the lumen of the tubules and collecting duct to undergo modification in the formation of the final product, urine.

Glomerular Filtration

The forces which control glomerular filtration are exactly the same as those that regulate the filtration of fluid from any capillary in the body, namely the resultant of hydrostatic and osmotic pressures. The hydrostatic pressure in the glomerular capillary is estimated to be about 65 mm Hg which is considerably higher than in other capillaries. The blood osmotic pressure is about 25 mm Hg. There is probably a back pressure, that is, hydrostatic pressure in the capsule and tubule of about 10 mm Hg. Since the filtrate is protein free, it exerts no osmotic pressure. Thus, the **effective glomerular filtration pressure** is calculated to be about 30 mm Hg [65 − (25 + 10)].

Samples of glomerular filtrate have been successfully collected in frogs and mammals. Analysis of these samples supports the fundamental concept that the glomerular filtrate is an ultrafiltrate of the plasma. In other words, the fluid forms simply by a process of filtration; there is no evidence of active secretion.

Tubular Reabsorption

After a part of the plasma has filtered through the glomerulas, the fluid passes through the lumen of the tubules, and there is a marked

modification of the filtrate. This can be understood when we consider that it has been shown that in man, about 130 ml of filtrate are formed per minute by both kidneys. However, since the normal rate of urine formation is only approximately 1 ml per minute, it is evident that the major part of the glomerular filtrate is reabsorbed by the blood as the fluid passes through the tubules. This is true because pressure in the capillary that surrounds the tubule is apparently much lower than it is in the glomerular capillary. As the blood in the tubular capillary has a very high osmotic pressure, water is thought to diffuse from the lumen of the tubule back into the blood. Furthermore, in the distal tubule, according to the classic concept, even more water is reabsorbed. This latter transfer is said to take place against concentration and pressure gradients, and therefore requires an active transport mechanism. The mechanism of this active transfer is not known, and whether or not there is such a mechanism is now seriously questioned (see counter-current hypothesis below).

All of the solutes which normally appear in urine are either completely, or in part, absorbed by the renal tubule. Much of the reabsorption of these substances represents passive diffusion in accord with the movement of water. But in addition, there is active transport of some substances. It is by virtue of the specific, independent transport mechanisms that the concentrations of the various components of the plasma are regulated.

The location of reabsorption of each substance varies. Some are primarily reabsorbed in the proximal tubule, some in the distal, and others reabsorbed in part by both. In the proximal tubules, glucose, phosphate, bicarbonate, potassium, sodium, the amino acids, and any protein that may have leaked into the filtrate are reabsorbed. In the distal tubule, water, chloride, and more sodium are reabsorbed.

Other substances, such as urea, uric acid, and sulfates, are partially reabsorbed by the renal tubule, but in lesser amounts than water. For example, about 99 percent of the filtered water is reabsorbed, but only about 50 per cent of urea is removed by the tubules. This suggests that such substances are not actively reabsorbed, but to the contrary, are only inadvertently reabsorbed with the transport of water. Most of the end-products of metabolism fall into this category.

Tubular Secretion

The renal tubules are capable of transporting substances from the blood into the filtrate, as well as from the filtrate back into the blood. This process is termed tubular secretion.

Most of the substances that the tubule handles in this manner are foreign to the body. This category includes diodrast, para-amino-hippuric acid, penicillin, and phenosulfonphthalein. There is also evidence that creatinine and potassium may be handled in this way. In other words, the concept is growing that some substances may move in two directions simultaneously through the tubule wall. Thus, it is postulated that potassium is reabsorbed from the filtrate, but it also is secreted into the filtrate. The final concentration depends upon the balance between these two opposing mechanisms.

The Counter-Current Hypothesis

Mention must be made of a relatively new concept concerning the formation of urine. It is called the counter-current hypothesis because the hair-pin bend of the nephron in combination with the collecting duct is visualized as constituting tubes in which the current, or direction of flow is in opposite, or counter, directions (Fig. 14.2).

It is postulated that there is exchange of water and some solutes between these opposing currents. Thus, it is maintained by proponents of this hypothesis that water moves out of the descending (proximal) tubule, crosses the tissue separating the tubules to enter the ascending (distal) tubule. There is said to be similar transfer from the collecting duct into the ascending tubule. Solutes, for example, sodium, are said to move in the opposite direction.

There is considerable evidence to support this hypothesis. The best evidence is obtained from studies in which samples are drawn from various parts of the nephron. It has been found that the maximum osmotic concentration is in urine taken from the thin loop, or from the collecting duct in the region of the thin loop. The osmotic concentration of the urine in the upper part of the ascending tubule is approximately the same as that in the upper part of the descending tubule. In such a system the degree of concentration is a function of the length of the tubules.

Further evidence for this hypothesis comes from studies in which the kidney is frozen and then sectioned. The osmolality is found to increase progressively in the slices ranging from the outside in toward the medulla. In other words, slices taken in the region of the

GLOMERULUS

DILUTE (ISOTONIC)

ASCENDING LIMB

COLLECTING DUCT

DESCENDING LIMB

H_2O

H_2O

Na

Na

H_2O

H_2O

Na

Na

H_2O

CONCENTRATED (HYPERTONIC)
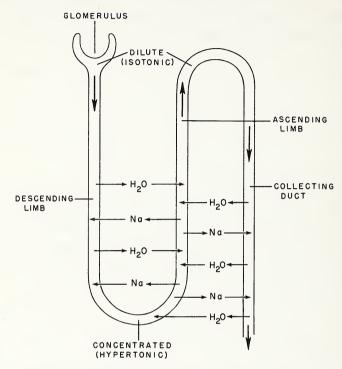
Fig. 14.2. Counter-current Hypothesis. By virtue of the reverse movements of water and solute the glomerular filtrate is progressively concentrated reaching a maximum in the loop of Henle. It then becomes progressively less concentrated as it ascends, and finally more concentrated in the collecting duct. The final concentration of the urine depends upon the concentration in the interstitial fluid surrounding the collecting duct.

thin loop have the highest osmolality. According to this hypothesis, then, the increasing concentration of solutes in the interstitial fluid causes water to move passively from the collecting duct. In short, it is believed that urine is concentrated by the passive removal of water in contradistinction to the classic concept that visualizes an active transport of water.

Composition of Urine

The composition of urine varies remarkably; the figures given in Table 14.1 are but averages. Renal function is normally capable of maintaining the constancy of the body fluids despite wide alteration in ingestion and in metabolic activity.

TABLE 14.1. Partial Composition of Urine Compared with Plasma

Substance	Plasma	Urine	24 hr. Urine
	mEq/l	mEq/l	grams
Sodium	140	120	4.0
Potassium	4.6	50	3.0
Calcium	4.5	6	0.2
Magnesium	1.9	9	0.1
Ammonia	—	40	1.1
Chloride	103	180	9.5
Bicarbonate	27	17	1.5
Sulfate	1.1	35	1.1
Phosphate	1.8	40	1.2
	mg%	mg%	grams
Glucose	80	0	0
Urea	2.5	1600	24
Uric acid	4	53	0.8
Creatinine	1.5	100	1.5
pH	7.41	6.0	6.0

CONTROL OF RENAL FUNCTION

It has been seen that the formation of urine is the result of two processes: 1) glomerular filtration, and 2) tubular modification. The adjustments, then, that alter the volume and composition of the urine so as to maintain the constancy of the internal environment must involve one or both of these processes.

Control of Glomerular Filtration

As previously mentioned, glomerular filtration depends upon the resultant between hydrostatic and osmotic pressures. The hydrostatic pressure in the glomerular capillary can be altered by:

1. *Arterial Blood Pressure.* A fall in arterial blood decreases urine flow primarily by decreasing the glomerular filtration rate due to lowered filtration pressure. When the mean arterial blood pressure decreases below about 60 mm Hg, urine production ceases.

It is usually taught that elevations in arterial blood pressure do not alter glomerular filtration pressure or renal blood flow. However, a recent study casts considerable doubt upon this conclusion; thus until more work is carried out it can only be concluded that under some conditions elevation of arterial blood pressure does increase glomerular filtration, and under other conditions it does not.

2. *Afferent and Efferent Arteriolar Resistance.* A change in resistance in the afferent and efferent arterioles can maintain normal renal function in the face of elevated arterial blood pressure. The arterioles are under autonomic nervous system control, and they also respond to various chemicals to regulate glomerular filtration. Activation of the sympathetic fibers that innervate the kidneys produces vasoconstriction of both arterioles. As a result, renal blood flow drops sharply, glomerular filtration ceases, and so does urine formation.

Epinephrine has an interesting influence on renal function. The over-all effect is to increase urine output and this has been explained on the basis that epinephrine causes greater constriction of the efferent arteriole than of the afferent arteriole, and therefore glomerular filtration pressure and rate are increased.

By virtue of arteriolar adjustments the kidney can usually maintain normal function despite widespread circulatory changes in the rest of the body. In addition, the kidney is an important organ in the regulation of circulation. In exercise, for example, renal blood flow decreases, thus making blood available for the active muscles. In circulatory shock, renal blood flow is decreased which aids the condition. At the same time urine production diminishes or ceases, and thus body water is conserved.

Control of Tubular Function

Although some substances do compete for the same tubular transport mechanism, in most cases these mechanisms are independent processes which can vary the excretion of specific substances.

1. *Water.* The excretion of water is under the control of the

antidiuretic hormone of the neurohypophysis. Under the influence of this hormone the reabsorption of water may be so great as to reduce the urine output to about 0.35 ml per minute. In the complete absence of the hormone, about 15-18 ml of urine per minute are excreted with a specific gravity practically the same as that of blood. The adrenocortical steroids and perhaps the gonadal steroids may also influence the reabsorption of water, and thus play roles in controlling urine volume.

2. *Sodium and Potassium*. The level of sodium and potassium in the body fluids is markedly influenced by the steroids secreted by the adrenal cortex. Part of this control is due to the action of these hormones on renal function. The adrenocortical steroids increase the reabsorption of sodium and decrease the reabsorption of potassium. Potassium is also actively secreted by the renal tubules, but what action the steroids have on this transport mechanism is not known.

3. *Calcium and Phosphate*. The tubular control of these ions is under the control of the hormone of the parathyroid gland. The parathyroid hormone increases calcium reabsorption and decreases the reabsorption of phosphate.

4. *Other Urinary Constituents*. Glucose and amino acids are very actively reabsorbed, but the rate of reabsorption seems to depend upon the filtered load up to a point where the reabsorptive mechanism becomes satiated. After this point is reached, the excess is excreted. This rate of reabsorption, in effect, provides a threshold that functions like a dam to maintain a specific level in the extracellular fluid. Glucose, for example, is completely reabsorbed by the tubules, thus none normally appears in the urine. But if the amount of glucose in the blood is increased high enough, then glucose will be excreted in the urine. The end-products of metabolism seem to be only passively reabsorbed and therefore as the filtered load increases the amount reabsorbed increases, but so does the quantity excreted.

In short, water and plasma ions have specific transport mechanisms which control their excretion so as to maintain concentrations in the body fluid within very narrow ranges. Other constituents increase in the urine as they tend to rise in the body fluids. In this way, undue concentrations in the body are prevented.

Diuretics

A diuretic is a substance that increases urine flow. This may be accomplished either by increasing glomerular filtration rate, or by decreasing the reabsorption of water. The more common diuretics include:

1. *Osmotic Diuretics.* Any substance that is freely filtered by the glomerulus, but is not readily absorbed by the tubule will cause diuresis if its concentration in the blood, and therefore in the filtrate, increases. Sucrose is a good example. When this substance appears in the filtrate, it impedes the reabsorption of water because it is not readily reabsorbed. Consequently, the osmotic pressure of the filtrate opposes the reabsorption of water.

2. *Xanthines and Alcohol.* These substances are thought to increase urine flow primarily by augmenting the rate of glomerular filtration. The mechanism is believed to be dilatation of the afferent arteriole which increases renal blood flow and elevates glomerular filtration pressure.

3. *Mercurial Compounds.* The mercurial compounds are commonly used as diuretics. Their action is probably at least twofold. In the first place, since they are not readily absorbed by the tubules they act osmotically. Secondly, they impede the reabsorption of sodium in the proximal tubule, which is probably the more important action. Because of the inhibition of sodium reabsorption, water reabsorption is decreased. However, exactly how mercury decreases sodium reabsorption is not known. It is thought to act on the responsible enzyme system.

MICTURITION

After the urine is formed in the individual nephrons, it flows through the collecting tubules into the renal pelvis. From the renal pelvis it moves down the ureters to the bladder. After a stay of a variable length of time in the bladder, it then flows through the urethra to be eliminated from the body. The elimination of urine is termed micturition, or simply, urination.

Ureters

The ureters are not merely passive ducts. They are equipped with smooth muscle which permits peristaltic movements. These move-

ments propel the fluid along. If one observes the urine entering the urinary bladder it is found to enter in irregular spurts. Likewise, if the drop pattern is recorded with a catheter in the lower part of the ureter, the drops are found to be irregularly spaced. In contradistinction, if the catheter is placed in the renal pelvis, the drop pattern is regular. This shows that urine formation is steady, but due to peristaltic movements of the ureters, the urine is moved along in spurts.

Anatomy of the Bladder

The urinary bladder is a hollow, pear-shaped muscular organ. The bladder musculature consists of smooth muscle fibers which collectively are referred to as the **detrusor muscle.** It is innervated by both sympathetic and parasympathetic fibers. The urethra and external sphincter possess a somatic, or voluntary, motor innervation.

Function of the Bladder

Evidence has been recently published which suggests that a passive diffusion of water and solutes can occur in the ureters and urinary bladder. Thus, as the highly concentrated urine accumulates in the bladder, substances that are more concentrated in the urine than they are in the blood, diffuse from the bladder to the blood. Conversely, water may move into the urine. It would therefore seem that the urine is modified after it leaves the kidney, but just how significant this modification is has not been clarified.

The urinary bladder does not behave simply as an elastic bag. As fluid collects in the bladder, the internal pressure increases very little. Were it otherwise, a back pressure would develop which could impede glomerular filtration.

The bladder wall contains stretch receptors. Impulses which they initiate are propagated via visceral afferent neurons in the pelvic nerve to the sacral cord. Ultimately, the parasymphathetic neurons are activated and, as a result, the detrusor muscle contracts (Fig. 14.3). This sharply increases the pressure within the bladder, the sphincters are forced open, and urine is eliminated.

Contraction of the detrusor muscle is but one reflex responsible for urination; another reflex controls the internal sphincter. This sphincter cannot be voluntarily opened or closed independently of de-

trusor contraction and relaxation. The external sphincter, on the other hand, can be vigorously closed despite detrusor contraction. Normally, a sequence of reflex actions occurs which causes detrusor contraction, and relaxation of both sphincter muscles. Following elimination, the sphincters close and the bladder relaxes.
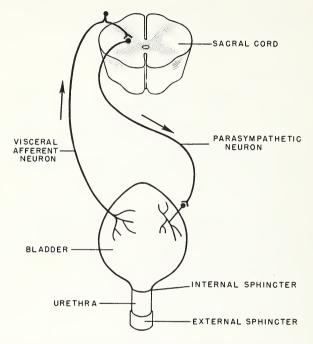
Fig. 14.3. Basic Reflex Controlling Urination. Distention of the bladder due to urine accumulation fires the reflex.

Higher Center Control

The reflexes that control the bladder are complete at the spinal cord level. After cord transection and the disappearance of spinal shock, the bladder still exhibits a micturition reflex. This shows that the reflex is complete at the spinal cord level, but it is also true that higher centers can influence this reflex.

The **anterior pons** gives rise to a descending tract which facilitates the micturition reflex. There is also a **midbrain** area which is inhibitory. It is not certain whether the inhibition is exerted on the pontine facilitatory center, or directly on the motor neurons that control the detrusor muscle. Finally, the cerebral cortex and **hypo-**

thalamus also play roles. They have an inhibitory effect. In short, there is the usual play of inhibitory and facilitatory impulses on the reflex arc. It is the resultant of these opposing influences that determines the sensitivity of the reflex response.

SUMMARY

The primary function of the kidneys is to regulate the volume and composition of the body fluids. The functional unit of each kidney is the nephron which consists of a **glomerulus, proximal tubule, thin loop of Henle,** and the **distal tube.** The capillary of the glomerulus lies between two arterioles: 1) the **afferent arteriole** and the **efferent arteriole.** Another capillary bed then serves the tubules and **collecting ducts.**

Urine is formed by the passive filtration of a filtrate of the plasma in the glomerulus. This filtrate is then modified as it passes through the remainder of the nephron and collecting duct. The **effective glomerular filtration pressure** determines the rate of filtration. In man about 130 ml of filtrate forms per minute. The normal rate of urine formation is only about 1 ml per minute. In the tubules and collecting ducts the filtrate is modified by the reabsorption of water and certain solutes, and the active secretion of other solutes. The final urine product is a hypertonic solution with a pH that averages about 6.0.

Glomerular filtration varies with the arterial blood pressure and with the resistance of the afferent and efferent arterioles. The volume of urine is controlled by the amount of **antidiuretic hormone** present. The **adrenal steroids** influence the excretion of sodium and potassium. The **parathyroid hormone** regulates calcium and phosphate excretion. Glucose does not normally appear in the urine, but does if the blood concentration increases above about 160 mg percent.

A **diuretic** is a substance that increases urine flow. Diuretics act either by increasing glomerular filtration, or by decreasing the reabsorption of water.

Urine is propelled in the **ureters** by peristaltic movements to enter and to accumulate in the **urinary bladder.** A stretch reflex causes contraction of the bladder and relaxation of the sphincters so that the urine is expelled from the bladder through the **urethra.** The **micturi-**

tion reflexes are modified by activity of the **anterior pons,** the **midbrain, hypothalamus,** and **cerebral cortex.**

Problems

1. Assume a glomerular filtration rate of 130 ml per minute and a urine flow of 1 ml per minute. What percentage change in water reabsorption is necessary to produce a 100 percent increase in urine flow?
2. In circulatory shock the arterial blood pressure falls below about 60 or 70 mm Hg. Under these conditions urine flow ceases. Why?
3. Epinephrine usually increases urine formation. This substance apparently causes greater constriction of the efferent than of the afferent arteriole. How would this increase urine formation?
4. How do the classical theory of urine formation and the counter-current hypothesis differ?
5. Individuals who have suffered transection of the spinal cord develop what is known as an autonomic bladder. This means that when the bladder accumulates a few hundred ml of urine it automatically empties. Explain the mechanism.

Additional Reading

Gottschalk, C. W. and Mylle, M., "Micropuncture Study of the Mammalian Urinary Concentrating Mechanism: Evidence for the Counter-Current Hypothesis," *Am. J. Physiol.,* **196,** 927-936 (1959).

Langley, L. L., "Outline of Physiology," McGraw-Hill Book Co., New York, N. Y., 1961.

Robinson, J. R., "Reflections on Renal Function," Charles C Thomas, Springfield, Ill., 1954.

Smith, H. W., "Principles of Renal Function," Oxford University Press, New York, N. Y., 1956.

THE ENDOCRINE SYSTEM

AN ENDOCRINE GLAND has already been defined as a gland that does not have a duct but rather secretes its product directly into the surrounding area or into the blood to be carried to various parts of the body where the product exerts its influence. The secretion of an endocrine gland is termed a **hormone**. Hormones regulate many metabolic processes; in this sense they are very similar to enzymes. And, as a matter of fact, as knowledge of hormones and enzymes increases, the difference between them lessens considerably.

The body is replete with hormones which regulate myriad functions. For example, acetylcholine may be considered to be a hormone secreted by the axon of a nerve to excite another nerve, or muscle. In the gastrointestinal tract there are many hormones that regulate the secretion of the various digestive juices and which also influence the smooth muscle that makes up the wall of the tract. Just about every cell in the body produces histamine. Although the precise role of this substance in the normal physiology is not clear, it is known that it increases capillary permeability and causes vasodilatation. And finally, there are specialized endocrine glands such as the thyroid, parathyroids, adrenals, gonads, and hypophysis which secrete a variety of hormones that have wide-spread and essential regulatory roles. This chapter will concern itself primarily with the function of these endocrine glands.

⸳ THYROID GLAND

The thyroid gland, located in the anterior part of the neck, weighs about 30 grams in man. The gland is arranged very much like a beehive. Each cubicle is termed a **follicle,** and the follicles normally contain a colloid. The thyroid has a copious blood supply which in man flows at an average rate of about 5 ml/min/gm of tissue.

Thyroglobulin

The colloid contained within the follicles of the thyroid gland is a protein which is organically combined with iodine. It is termed thyroglobulin. It has a molecular weight of about 675,000 and contains approximately 0.75 percent iodine. Thyroglobulin, under the influence of a proteolytic enzyme releases iodothyronines which are secreted by the gland.

Thyroid Hormones

The thyroid gland is now considered to secrete at least four hormones. All of them are **iodothyronines** (Fig. 15.1). Of the four hormones, 3,5,3'-triiodothyronine is the most active. It is believed to be from 5 to 10% more potent than 3,5,3'5'-tetraiodothyronine (thyroxine). But the rate of thyroxine formation is by far the greater.

Quite clearly, iodine is an essential component of the thyroid hormones. Iodine is avidly accumulated by the thyroid gland and then incorporated into the hormones.

There is no question but that the thyroid hormones strikingly increase metabolic activity in perhaps every cell of the body. How this is accomplished is not completely understood. It has been suggested that the thyroid hormones stimulate enzyme production. The greater the concentration of a specific enzyme, the faster will that reaction proceed. If this hypothesis proves correct, then the very broad and generalized influence of the hormones on metabolism becomes explicable. More specifically, it has been fairly well demonstrated, at least *in vitro*, that thyroid hormones have an uncoupling effect on oxidative phosphorylation. Although the exact details have still to be clarified, it is believed that oxidation and phosphorylation are "coupled." This means that the reactions for both processes are interrelated so that the rate at which foodstuffs are burned is in step

3,5,3',5' – TETRAIODOTHYRONINE

3,5,3' – TRIIODOTHYRONINE

3,3',5' – TRIIODOTHYRONINE

3,3' – DIIODOTHYRONINE

Fig. 15.1.   Iodothyronines Secreted by the Thyroid Gland.

with the energy requirements of the cell. When oxidation and phosphorylation are uncoupled, some of the energy from cellular oxidation is liberated in the form of heat, whereas if uncoupling had not occurred that energy would have been used for synthesis of high-energy phosphorus compounds. If thyroxine truly has this uncoupling action *in vivo,* the well documented influence of the hormone on the metabolic rate becomes explicable.

Regulation of Thyroid Function

The thyroid gland is regulated by the **thyrotropic hormone** secreted by the anterior lobe of the hypophyseal gland. The thyrotropic hormone, also called thyroid stimulating hormone, **TSH,** acts on the thyroid gland to increase its activity and thereby elevate the level of circulating thyroid hormones.
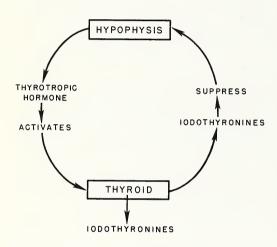
Fig. 15.2. Regulation of Thyroid Activity.

There is a reciprocal regulation between TSH and the thyroid hormones (Fig. 15.2). This means that when the blood level concentration of the thyroid hormones decreases, more TSH is secreted. The TSH then causes more iodothyronine to be secreted by the thyroid gland. A balance is quickly reached, however, so that there is a steady concentration of TSH and of the thyroid hormones in the blood. This reciprocal, or feedback mechanism, functions most effectively. Thus, if excessive thyroid hormone, for any reason, is utilized, the blood level will fall, more TSH will be liberated, and the output of the thyroid gland will increase.

The thyroid hormones undergo deiodination, or removal of the amino group, in the liver and perhaps in most other tissues as well. What relation these metabolic processes have to the function of the hormones is not known. But the rate at which the hormones are

metabolized, as just explained, plays an important role in the blood level. In short, by this self-regulating mechanism ample thyroid hormone is always available for metabolic needs.

PARATHYROID GLANDS

The parathyroid glands, usually four in number, lie in intimate association with the thyroid gland. The total parathyroid tissue weighs only about 0.1 gram.

Parathyroid Hormone

Although it has been known since 1925 that the parathyroid glands secrete a hormone, the chemical nature of the substance is still not known with certainty. The hormone appears to be a polypeptide containing 76 amino acid units.

It has been well demonstrated that the parathyroid hormone enhances the absorption of calcium by the gut. Phosphate, apparently, is not influenced. When the hormone is administered over a period of time, a positive calcium balance results due, in part, to increased gut absorption.

One of the major effects of the parathyroid hormone is to cause demineralization of bone. The hormone increases both the number and the activity of the osteoclasts and, in this way, causes dissolution of bone.

It is now well documented that the parathyroid hormone influences the renal handling of phosphate. Administration of the hormone increases excretion of phosphate. At the same time the excretion of calcium is decreased. The renal tubular transport mechanisms for phosphate and calcium are not known, but it is clear that whatever their nature, they are regulated by the parathyroid hormone.

By virtue of the influence of parathyroid hormone on the gut, bone, and the kidneys, calcium and phosphate metabolism is regulated. Thus, the hormone causes an elevation in serum calcium and a decrease in phosphate. A shortage of hormone results in a lowered serum calcium and an elevated serum phosphate.

Regulation of Parathyroid Function

There is general agreement that the level of serum calcium regulates parathyroid function because a decrease in calcium level in the

blood, no matter how produced, results in increased parathyroid activity. More hormone is secreted and this, as previously indicated, elevates the serum calcium level. As a result, the parathyroid glands are suppressed (Fig. 15.3). There is thus a self-regulating mechanism for serum calcium.
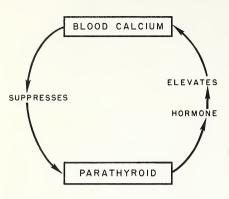
Fig. 15.3.   Regulation of Parathyroid Activity.

THE PANCREAS

The pancreas has a dual role; it is both an exocrine and an endocrine organ. Its exocrine function is to secrete pancreatic juice which is carried by the pancreatic duct to the duodenum for digestive purposes. Its endocrine function is to elaborate two hormones which regulate carbohydrate metabolism. The acinar cells are responsible for the secretion of the pancreatic juice. The beta cells of the islets of Langerhans produce one hormone, termed **insulin.** The alpha cells of the islets of Langerhans produce the other hormone, termed **glucagon.**

Hypoglycemic Hormone

Insulin causes blood sugar to decrease and therefore it is termed the hypoglycemic hormone. Insulin is a polypeptide with a molecular weight of about 6,000.

Function of Insulin. Insulin regulates carbohydrate metabolism. Injection of the hormone lowers blood sugar by increasing the rate of carbohydrate utilization in the cells, and by enhancing liver and muscle glycogen formation. It is now believed that the primary role of insulin is to facilitate the entry of glucose and perhaps other monosaccharides into the cell.

Insulin also influences fat metabolism. If there are inadequate quantities of insulin, **ketosis** results. Ketosis means excessive ketone bodies in the blood and urine. This result is thought to be secondary to disturbed carbohydrate metabolism. In other words, in the

absence of insulin, carbohydrate is not properly utilized and there-fore fat must undergo excessive catabolism with the formation of ketone bodies. For the same reason there is excessive catabolism of protein and a consequent elevation in blood amino acids. Because protein anabolism is essential to growth, it is found that growth is retarded in the absence of insulin.

Regulation of Insulin Secretion. The secretion of insulin by the islet cells is apparently controlled by the level of the blood sugar. Elevation of blood sugar suppresses insulin secretion; low blood sugar augments it. There is thus a feedback mechanism by which blood sugar is maintained within a narrow range.

Hyperglycemic Hormone

The hormone of the pancreas that elevates blood sugar is termed glucagon. It is thought that this augmented glycogenolysis results from an influence the hormone has on liver phosphorylase phospho-kinase.

The regulation of glucagon secretion and its physiological role have not yet been clarified.

ADRENAL GLANDS

The adrenal glands in man are considered to be at least two dis-tinct endocrine glands in one. The inner core is termed the **adrenal medulla,** and the outer rim is called the **adrenal cortex.** The adrenal cortex may be further subdivided into three layers, each of which elaborates at least one hormone. These layers are termed from the outermost inward, zona glomerulosa, zona fasciculata, and zona reticularis.

Each adrenal gland, in man, weighs about 6 grams. They have a very copious blood supply which averages from 6 to 7 ml per minute per gram.

The Adrenal Medulla

It is possible to remove the adrenal medulla without disrupting the function of the adrenal cortex. When this is done the animal quickly recovers and appears to be quite normal. But this does not

mean that the medullary tissue has no function. In the first place, there is chromaffin tissue throughout the body and therefore removal of the adrenal medulla does not completely eradicate the source of the hormones. In the second place, the medullary hormones have as their primary function, the maintenance of the constancy of the internal environment in the face of adverse conditions. Thus, unless the animal is placed under stress, the results of hormone lack are difficult to observe.

Medullary Hormones. It is now known that the adrenal medulla secretes two hormones: 1) epinephrine, and 2) norepinephrine (Fig. 15.4). The latter is sometimes referred to as arterenol.

EPINEPHRINE   NOREPINEPHRINE

Fig. 15.4. The Adrenal Medullary Hormones.

The medullary hormones have an effect on smooth muscle, cardiac muscle, and on carbohydrate metabolism. Since smooth muscle has so many functions, it is not surprising that the hormones have such a widespread influence. In short, the medullary hormones elevate blood pressure, increase heart rate, augment the force of ventricular contraction, relax the smooth muscle of the gastrointestinal tract, inhibit peristalsis, dilate the bronchial musculature, enlarge the pupil, inhibit the urinary bladder, and contract the gallbladder.

The main role of epinephrine on carbohydrate metabolism is to elevate blood sugar by enhancing glycogenolysis both in the liver and in muscle. The specific effect appears to be to augment the action of the enzyme phosphorylase phosphokinase. In this respect the action is identical with glucagon.

The effects of the two medullary hormones are not always iden-

tical. For example, epinephrine increases heart rate, norepinephrine does not. Epinephrine, instead of constricting the arterioles of skeletal muscle, dilates them. Thus, if the total peripheral resistance is determined, it may be found that norepinephrine increases it while epinephrine evokes a decrease. Likewise, epinephrine has a much greater influence on carbohydrate metabolism than does norepinephrine.

Regulation of Medullary Hormone Secretion. The two hormones are apparently stored in, and secreted from, separate cells in the medulla. It has also been shown that they are secreted in varying quantities in response to different stimuli. And since they do vary in their effects, the precise mechanism which normally controls the elaboration of each hormone remains to be clarified.

The medullary tissue is innervated by sympathetic neurons. Activation of these neurons results in the secretion of both hormones. Yet, it has been shown that stimulation of different nuclei of the hypothalamus causes the medullary tissue to secrete the hormones in varying proportions. This would seem to indicate that the specific cells responsible for each hormone have a distinct and separate innervation.

The Adrenal Cortex

If both adrenal glands are removed, the untreated animal dies in' about 7 days. Death results because of the loss of the adrenal cortex, not the medulla.

Cortical Hormones. A crude extract of the adrenal cortex contains many distinct compounds, all of which are steroids. Some 30 steroids have now been identified. The human adrenal gland appears to secrete at least the following steroids in significant amounts: 1) 17-alpha-hydroxycorticosterone also termed **cortisol** as well as **hydrocortisone** (Fig. 15.5), 2) 11-dehydro-17 alpha-hydroxycorticosterone which is more conveniently called **cortisone,** 3) **aldosterone**, and 4) **androgens.** It is believed that aldosterone is formed in the zona glomerulosa, cortisol and cortisone in the zona fasciculata, and androgens in the zona reticularis.

The cortical steroids evoke many alterations. In general, the two most important influences exerted by these hormones are on carbo-

hydrate and electrolyte metabolism. Some of the steroids evoke carbohydrate changes, whereas others affect electrolytes. For this reason it has become customary to speak of **glucocorticoids** and **mineralocorticoids.** It is true, however, that the glucocorticoids do have some influence on electrolytes, and the mineralocorticoids can alter carbohydrate metabolism.

CORTISOL

CORTISONE

ALDOSTERONE

ADRENOSTERONE

Fig. 15.5. The Adrenal Cortical Hormones.

Cortisol and cortisone cause hyperglycemia. These steroids hasten the conversion of protein to carbohydrate. They are thought to inhibit the synthesis of protein from amino acids, thus keeping a high supply of amino acids available for conversion to glycogen. Whether they also then hasten the conversion of glycogen to blood sugar is not clear, although certainly blood sugar does increase.

Administration of cortisol, or cortisone, results in excessive fat

catabolism with the development of ketosis. Again, these changes are thought to reflect derangements in carbohydrate metabolism. Apparently the steroids not only enhance gluconeogenesis, but also inhibit the utilization of glucose. For energy needs, the organism uses fat.

The electrolytes, sodium and potassium, are markedly influenced by the adrenal steroids, especially by aldosterone. This hormone acts on the renal tubules to increase the active reabsorption of sodium. There is also greater absorption of water. As a result, under the influence of this hormone sodium and water accumulate in the body. Conversely, aldosterone increases the concentration of potassium in the urine. In addition to the effect of aldosterone on renal function it also acts on most cells of the body to regulate the movement of sodium and potassium; thus, potassium moves into the cell and sodium moves out. As a result of the influence of aldosterone on renal function and on the cells there is an elevation of plasma sodium and a decrease of potassium. Changes in calcium metabolism have also been reported. Cortical steroid administration leads to augmented excretion of calcium. But this is thought to be secondary to inhibition of protein anabolism.

The steroids secreted by the zona reticularis appear to be identical to those elaborated by the gonads. The major hormone of this part of the adrenal cortex is similar to the male sex hormone and is therefore termed androgen, but under certain circumstances female sex hormones may be secreted.

Regulation of Cortical Steroid Secretion. The anterior lobe of the hypophyseal gland secretes a hormone, termed the **adrenocorticotropic hormone**, ACTH, which regulates the zona fasciculata and which can also influence the zona glomerulosa. However, what influence it has on the zona reticularis is not clear. There is a reciprocal relationship between cortisol and cortisone on the one hand, and ACTH on the other. Thus, the levels of cortisol and cortisone are self-regulating.

The control of the zona glomerulosa and the secretion of aldosterone is still a very open problem. ACTH, as mentioned, can alter the activity of this part of the cortex, but there are reasons to believe that there are other, more important regulatory forces. The concen-

tration of sodium or potassium in the circulating blood may directly control the output of aldosterone.

THE GONADS

The term gonad means "seed." It refers to one of the functions of the gonads, namely to produce the sexual seed. But the gonads also have another function, and that is to elaborate hormones which are essential to sexual activity and to reproduction. The gonads in the male are termed the **testes;** whereas, in the female they are referred to as the **ovaries.**

The Ovaries

The ovaries, two in number, lie deep in the pelvic cavity, one on either side of the uterus. They are small, roughly spherical bodies which weigh about 6 grams each.

Female Sex Hormones. There are two types of female sex hormones: 1) estrogens and 2) progesterone. In the human female, at least three different estrogens have been identified: 1) estradiol, 2) estrone, and 3) estriol (Fig. 15.6). It is believed that the ovaries secrete estradiol which is then converted into estrone, then into estriol, and finally into non-estrogenic steroids during the process of metabolism.

The estrogens are responsible for the growth and development of the female sexual organs, the secondary sexual characteristics, and for sexual drive. The estrogens also influence bone growth; in fact, the characteristic spurt in growth immediately following puberty is due to the influence of these hormones on the long bones of the body. But the estrogens also cause closure of the epiphyseal centers so bone development ceases earlier than it does when the ovaries are not present, or when they mature late. In addition, the estrogens influence fat metabolism so as to cause fat deposition in the subcutaneous tissues. Fat, in the female, is characteristically deposited in the hips, breasts, and thighs.

Progesterone is often referred to as the hormone of pregnancy. Its function is certainly to prepare the uterus for implantation of the fertilized ovum and for maintaining it during pregnancy. Progesterone also helps to prepare the breasts for their function of lacta-

tion. The estrogens develop the duct system of the breast; whereas, progesterone is responsible for the development of the alveoli, causing the alveolar cells to become secretory.

Regulation of Ovarian Function. The ovaries are under the control of three hormones secreted by the anterior lobe of the hypophysis. Collectively they are termed **gonadotropic hormones,** or simply **gonadotropins.** They are all thought to be protein. One gonadotropin, the follicle-stimulating hormone, regulates development of the ovum. Another, the luteinizing hormone, is responsible for the formation of the corpus luteum. Luteotropic hormone, the third gonadotropin, is thought to stimulate the corpus luteum to secrete progesterone.

ESTRADIOL

ESTRONE

ESTRIOL

PROGESTERONE

Fig. 15.6.   The Ovarian Hormones.

There are complex reciprocal relationships between the ovarian and gonadotropic hormones. But rather than achieve a steady state as the thyroid and adrenocortical hormones do, the female sex hormones and the gonadotropins rise and fall in a monthly pattern and are thus responsible for the **menstrual cycle.**

The Testes

In the developing embryo, the testes and the ovaries arise from the same tissue and form at the same site. However, before birth, at about the sixth or seventh month, the testes normally descend through the inguinal canal to come to rest in the scrotal sac, where they reside throughout life.

Male Sex Hormones. All male hormones are classified as androgens. The testes secrete two types: 1) testosterone, and 2) androsterone (Fig. 15.7). Testosterone is at least 10 times more potent than androsterone and is secreted in far greater quantities, therefore it is often considered to be the only androgen of importance.

TESTOSTERONE   ANDROSTERONE

Fig. 15.7. The Testicular Hormones.

The androgens are responsible for the development of the sexual organs and the secondary sexual characteristics. It appears likely that the testes elaborate testosterone during fetal life. This androgen offsets the influence of maternal female sex hormones and is responsible for the development of the male characteristics during fetal life. It causes the descent of the testes into the scrotum before birth.

At puberty testosterone secretion increases and is responsible for widespread alterations. These include: sexual maturation, lowering of the voice, growth of the beard and body hair, changes in scalp hair, bone growth and early closure of the epiphyseal centers, and salt and water retention.

Regulation of Testicular Function. The anterior lobe of the hypophysis in the male, just as in the female, secretes gonadotropic hormones, but probably only two in number. These gonadotropins enter into a reciprocal relationship with the androgens. It is thought

that there is a waxing and waning of the androgens just as there is of the female sex hormones, but the significance of this cyle in the male is not clear.

HYPOPHYSEAL GLAND

The location of the hypophyseal gland at the base of the brain and the fact that it develops in conjunction with the oral cavity gave rise to the opinion among early investigators that this organ is concerned with the production of phlegm. Because of this erroneous concept, the structure was originally called the **pituitary** gland, and that name has persisted to this day. But since the term, hypophysis, means "to grow under," it is more accurate and therefore to be preferred.

The hypophyseal gland is a small, rounded body attached to the base of the brain by the hypophyseal stalk. In the adult it weighs about 0.6 gram. The gland is divided into an **anterior** and a **posterior** lobe. Because of its glandular and great secretory capacity, the anterior lobe is often referred to as the **adenohypophysis.** And since the posterior lobe is so richly innervated, it is called the **neurohypophysis.**

The Neurohypophysis

Most authorities now believe that the hormones of the posterior lobe are actually secreted by nuclei of the hypothalamus. These compounds are then thought to pass along the nerve fibers to enter the posterior lobe where they are stored until released into the circulation.

Neurohypophyseal Hormones. From the crude extract of the neurohypophysis at least 2 purified hormones can be obtained: 1) vasopressin, also known as the antidiuretic hormone, ADH; and 2) oxytocin.

Vasopressin, as the term indicates, is capable of elevating arterial blood pressure, but its more important function is to enhance the reabsorption of water by the kidneys. It will be recalled that in man about 130 ml of filtrate are formed by the glomeruli each minute. Approximately 129 ml of this filtrate are reabsorbed. This reabsorption, at least in part, is under the control of ADH.

Oxytocin means "rapid birth"; it acts on the uterus to evoke

powerful contractions and is often used to assist in childbirth. However, parturition appears to occur quite normally in the absence of the neurohypophysis and thus there is considerable doubt as to the importance of oxytocin in childbirth. On the other hand it is quite clear that oxytocin does play a role in lactation. However, whether the ejection of milk in response to suckling is due to a direct action of oxytocin on the breast, or whether oxytocin exerts this end-result by causing the release of another hormone, is still not certain.

Regulation of Neurohypophyseal Secretion. In the hypothalamus there are cells that seem to be sensitive to the osmotic pressure of the blood; they are, accordingly, termed **osmoreceptors.** During dehydration the osmotic pressure increases, the osmoreceptors fire and, as a result, impulses are propagated to the neurohypophysis to cause the release of ADH. ADH causes water retention which relieves the dehydrated state. There is thus a self-regulating feedback mechanism to control the body's hydration.

The release of oxytocin is brought about by suckling. The act of suckling initiates impulses which are propagated to the hypothalamus causing oxytocin secretion and release.

The Adenohypophysis

By virtue of the many hormones that the anterior lobe of the hypophysis secretes, many other endocrine glands are regulated. In this way, the adenohypophysis controls, or integrates, many physiological activities (Fig. 15.8). It will be recalled that the nervous system is a major regulatory system; the endocrine glands constitute another. It is now known that there is an intimate relationship between the nervous system and the endocrine glands exerted by **neurohumors** which modify adenohypophyseal function. Accordingly, these two great regulatory systems are capable of working together and in great harmony.

Adenohypophyseal Hormones. The adenohypophysis secretes at least six hormones: thyrotropin, adrenocorticotropin, growth hormone, and three gonadotropins.

The thyrotropic and adrenocorticotropic hormones, as already noted, control the thyroid gland and the adrenal cortices. The growth hormone, also called the somatotropic hormone, or simply **somatotropin,** plays an outstanding role in determining the growth
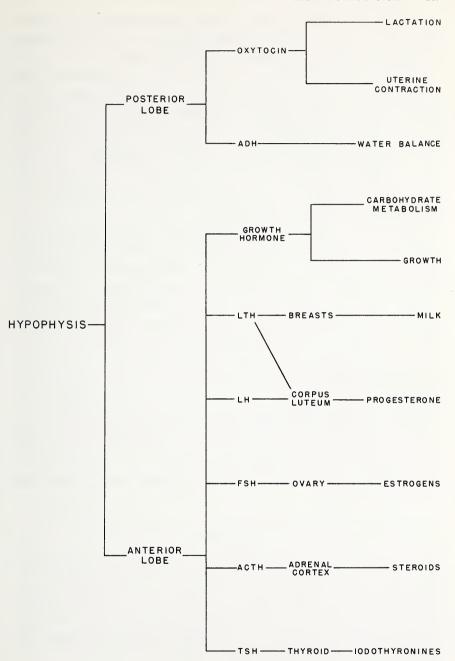
Fig. 15.8. Function of the Hypophyseal Hormones.

of the organism. This is a direct effect of the hormone. Unlike the other hormones secreted by the adenohypophysis, growth hormone does not function through another endocrine gland.

If somatotropin is given to an animal before the epiphyseal centers close, there results great acceleration and final growth of the bones as well as other tissues. In this process, growth hormone influences protein, carbohydrate, and lipid metabolism. It should be noted that growth hormone elevates blood sugar. Thus, it has been seen that insulin lowers blood sugar, but epinephrine, glucagon, corticosteroids, and growth hormone all elevate it. A balance between these hormones, then, is responsible for the maintenance of the blood sugar within a narrow range.

Regulation of Adenohypophyseal Secretion. The secretions of all the adenohypophyseal hormones, with the exception of somatotropin, are regulated, in part, by a reciprocal relationship with the hormones that they regulate. Thus, thyroid hormones decrease the secretion of TSH, adrenocortical hormones decease the secretion of ACTH, and the sex hormones decrease the secretion of the gonadotropins. But in addition to this control, there are neurohumors secreted by nuclei of the hypothalamus which regulate the hormonal output of the adenohypophysis. In other words, there is a self-regulating mechanism, but the "setting" of this mechanism can be altered by the nervous system acting through the hypothalamus. It is similar to a thermostat that keeps the temperature of a building constant. The setting on the thermostat can be changed and then the temperature is maintained at a different level.

Very little is known concerning the regulation of the secretion of the growth hormone.

SUMMARY

Endocrine, or ductless, glands secrete **hormones** which regulate cell physiology. The **thyroid** gland manufactures a thyroglobulin combined with iodine. This colloid then releases various **iodothyronines** which are the active thyroid hormones. They increase metabolic processes. The thyroid gland is controlled by the **thyrotropic hormone** secreted by the hypophysis.

The **parathyroid** glands secrete a polypeptide hormone that regulates calcium and phosphate metabolism by its action on bone and on the kidneys. The level of serum calcium controls the activity of the parathyroids.

The **pancreas** secretes a hypoglycemic hormone, **insulin**, and a hyperglycemic hormone, **glucagon**. Insulin regulates carbohydrate metabolism by increasing the rate of utilization; glucagon enhances the conversion of liver glycogen to glucose. The level of blood sugar controls the secretion of insulin. It is not known how glucagon is regulated.

The **adrenal** glands consist of the adrenal **medulla** and the adrenal **cortex**. The medulla secretes **epinephrine** and **norepinephrine**. These hormones influence smooth muscle, cardiac muscle, and carbohydrate metabolism. Their secretion is regulated by the sympathetic nervous system. The cortex secretes steroids which control carbohydrate and protein metabolism, and sodium and potassium metabolism. The **glucocorticoids** enhance the conversion of protein to carbohydrate and regulate the movement of potassium and sodium between the intra- and extracellular fluids. The **mineralocorticoids** act on the kidneys to increase potassium excretion and decrease sodium elimination. The adrenal cortex is regulated by the **adrenocorticotropic** hormone.

The **ovaries** secrete **estrogens** and **progesterone** which are responsible for sexual maturation, control of the menstrual cycle, and for development of the breasts. The ovaries are regulated by the **gonadotropins**. The **testes** secrete **testosterone** and **androsterone** which bring about sexual maturation. The testes are also regulated by the gonadotropins.

The **adenohypophysis** secretes tropic hormones and also the **growth hormone** which regulates growth of the organism and influences carbohydrate metabolism. The neurohypophysis is under the control of neurohumors secreted by the hypothalamus. The **neurohypophysis** releases **vasopressin**, also called the **antidiuretic hormone**, and **oxytocin**. The main role of ADH is to regulate water balance by its action on the kidneys. Oxytocin is essential for lactation and may assist childbirth. These hormones are formed in the hypothalamus and are under nervous system control.

Problems

1. Outline the major primary and secondary changes that would result from the removal of the hypophyseal gland.
2. If one adrenal gland is removed, the cortex of the remaining gland becomes enlarged and secretes a quantity of hormone about equal to that formerly secreted by the two glands combined. What is the mechanism responsible for this increased activity?
3. If the pregnant woman is not provided with an ample supply of milk her parathyroid glands enlarge. Why?
4. Discuss the hormonal control of carbohydrate metabolism.
5. An individual exercises vigorously and for an extended period of time in a hot, humid climate. Thereafter his urine output falls below the normal range. What is the mechanism responsible for this occurrence?
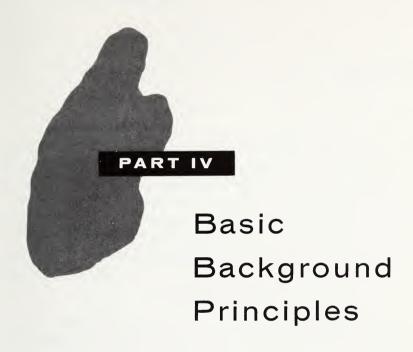
Additional Reading

Best, C. H., and Taylor, N. B., "The Physiological Basis of Medical Practice," 7th ed., The Williams & Wilkins Company, Baltimore, Md., 1961.

Fieser, L. F., and Fieser, M., "Steroids," Reinhold Publishing Corporation, New York, N. Y., 1959.

Gardiner-Hill, H., "Modern Trends in Endocrinology," Hoeber, New York, N. Y., 1956.

Hall, D. F., "The Function of the Endocrine Glands," W. B. Saunders Co., Philadelphia, Pa., 1959.

Ruch, T. C., and Fulton, J. F., "Medical Physiology and Biophysics," W. B. Saunders Co., Philadelphia, Pa., 1960.

Turner, C. D., "General Endocrinology," 3rd ed., W. B. Saunders Co., Philadelphia, Pa., 1960.

Velado, J. T., ed., "The Endocrinology of Reproduction," Oxford University Press, New York, N. Y., 1958.

PART IV

Basic
Background
Principles

CHAPTER 16

DATA

DATA ARE THE fundamental building blocks of all sciences. For the inept investigator instead of building blocks, data become stumbling blocks. There are many precise and involved statistical methods for evaluating data as we shall see in this chapter, but if the data is obtained inaccurately, no statistical method will be of value and neither will the data. Thus, before one can even think of formulating a hypothesis a consideration must be given to the methods of obtaining and analyzing data.

THE RAW DATA

Any investigation begins with a question, a problem. Generally the investigator is thoroughly familiar with his field of interest, and he is conversant with its total body of knowledge. Since this body has limits, peripheries, he desires to extend those limits, to press back the peripheries. It is as though he were travelling on a road that ends at the edge of the forest; he wants to see beyond. He desires to clear the forest and extend the road.

For the sake of simplicity let us select a relatively simple problem. Assume that it is desirable to know the mass, that is, the weight of the eggs of a particular species of frog.

Experimental Design

Once the problem is decided upon the question of experimental design must be faced. At this stage inestimable time and effort can be saved, and confusion avoided, if one thinks his problem through, and carefully outlines his procedure.

How many animals will be necessary? How will the eggs be han-

dled and prepared for observation so as to avoid alteration of the cell morphology? How will the cell weight be determined? How many cells from each animal will be weighed? These are the important considerations in the experimental design of the problem we have selected. As has already been emphasized, if the data are obtained inaccurately, they are meaningless; worse, like a poorly constructed signpost they may not only provide no information, they may send you in the wrong direction.

Measurement

As soon as several eggs have been weighed, it will be noted that the same value has not been obtained for each cell. Do the cells really differ? Undoubtedly. But do the differences in the obtained values reflect something other than cell variation? Yes. That something is known as **error.**

There are two types of error, **systematic** and **random.** As implied in the term, systematic error is caused by a defect in the system used to make the measurement. For example, a torsion balance may be used to weigh the cells. If that balance is not accurate then, of course, every reading will contain an error of approximately the same magnitude. If this error is very small in relation to the cell diameter, it may be ignored.

Random error represents the difference between the observed value and the true measurement caused by a variety of factors other than those inherent in the measuring system. With any measuring device there are limits to the degree to which the value may be read. Thus, in weighing the same cell several times the weight may be read a fraction of a unit too high or a fraction too low. But clearly, if several readings, or measurements, are made such errors will balance out, will cancel one another.

Two other terms are used in association with systematic and random error: **accuracy** and **precision.** No matter how many times a measurement is repeated the systematic error persists; it cannot be eliminated except by refining the measuring system. The term accuracy indicates the magnitude of the systematic error. If this error is great in relation to the cell weight, then the accuracy of the method is intolerably low.

If one cell is weighed several times and then the values averaged,

a figure will be obtained which should be relatively free of random error. If the same cell is once again measured and the values averaged, this second average will be identical with, or very close to, the first. This can be done many times, that is to say, the value is reproducible. Precision is a measure of this reproducibility. It should then be clear that our method may be very precise, made so by repetitive measurements, but it still may be intolerably inaccurate due to a fault in the measuring system.

Recording

The data may be kept in one's head, it may be jotted down on scraps of paper, or in a notebook, or on individual record sheets, or by the use of elaborate instruments it may be printed or coded on cards. Clearly, a complete and permanent record should be made. The design of that record varies with the experiment. Neatness and completeness are essential.

Let us assume that in the frog egg problem preliminary checking discloses the accuracy of the torsion balance used to weigh the cells to be very high. It is found further that any one cell may be weighed time after time with adequate precision. The frogs are all kept in the same controlled environment and a procedure worked out which permits collecting and handling 8-10 eggs at one time from each animal used. Work now begins and the raw data are collected. Such data may be easily entered directly into a notebook, or onto printed or mimeographed sheets. The design may be as follows:

DATE _____	FROG NUMBER __	WEIGHT ___ (grams)
	Egg No.	Weight (mg)
	1	7.0
	2	5.6
	3	5.5
	4	4.1
	5	5.5
	6	4.2
	7	5.6
	8	5.4
	9	5.5
	10	6.6
	Average 5.5	

The above is the simplest possible form designed to provide a permanent record of all the essential data. Space is sometimes left to be used for brief comments of observations which may or may not be pertinent to the present experiment.

The work continues until a group of eggs from several frogs of the same species have been weighed. How many? This is an important question. Every student realizes that more than one animal must be used for any experiment, but very few students understand how the actual number is determined. Beyond question the greater the number of animals used the greater the significance of the results, but there is a limit beyond which it is meaningless to proceed. That limit, as well as other pertinent information, is determined from the obtained data by a science called biostatistics.

BIOSTATISTICS

Statistics is defined as the science of collection and classification of facts on the basis of relative number or occurrence. It is the systematic compilation of instances for the inference of general truths. Biostatistics, then, is simply a science in which the mathematical principles of handling frequency distributions is applied to biological data.

The Mean

The mean of a group of data is the arithmetic average. In the sample data sheet presented above, the mean, or average, of the weights of 10 eggs from one frog was obtained by dividing the sum of the weights by the number of observations. Since there are certain mathematical symbols generally used in statistical work it is advisable to introduce these symbols at the very beginning. Thus:

$$x = \text{a single measurement}$$
$$\Sigma = \text{sum of}$$
$$n = \text{number of measurements}$$
$$\bar{x} = \text{the mean}$$

$$\text{Therefore, } \bar{x} = \frac{\Sigma x}{n}$$

Using the data from frog number 1:

$$\overline{x} = \frac{55}{10} = 5.5 \text{ mg}$$

The Range

Ten eggs from frog 1 were weighed. Inspection of the data discloses that the lightest egg weighed 4.1 mg and the heaviest 7.0. In this particular group, then, the range of weights is from 4.1 to 7.0 with a mean of 5.5. The range gives but a very rough idea of the variation of the data from the mean.

Standard Deviation

The standard deviation is widely used as an indication of the variation of the data around the mean. To determine this value the following symbols are used:

$$s = \text{standard deviation}$$
$$x = \text{a single measurement}$$
$$\overline{x} = \text{the mean}$$
$$n = \text{the number of measurements}$$
$$\Sigma = \text{sum of}$$

$$s = \sqrt{\frac{\Sigma (x - \overline{x})^2}{n}}$$

Without going into the derivation of this relationship, a moment's study will disclose that the equation considers two important factors: 1) the variation of each measurement from the mean, that is $x - \overline{x}$, and 2) the number of measurements, n. The standard deviation, which is an indication of variation, obviously will increase if $x - \overline{x}$ increases without a proportional increase in n, and, conversely, will decrease with a greater number of measurements unless $x - \overline{x}$ grows proportionally.

The more mathematically minded will immediately realize that the process of squaring and then extracting the square root is utilized simply to obviate the handling of values of different signs, that is, plus and minus. Squaring converts all of the minus values to plus.

In the frog-egg example only ten measurements were made, but even with this small number a pattern can be discerned:

Egg Weight (mg)	Number of Eggs of Same Weight
4.1	1
4.2	1
5.4	1
5.5	3
5.6	2
6.6	1
7.0	1

When these values are plotted as has been done in Fig. 16.1, a rather irregular curve results. But if instead of 10 eggs, 100 are weighed then the curve becomes smoother and more meaningful. When a very large number of measurements is made, the so-called **normal frequency distribution** curve evolves (Fig. 16.2).

The standard deviation in the frog-egg example is calculated as follows:

x	$\bar{x}$	$x - \bar{x}$	$(x - \bar{x})^2$
7.0	5.5	1.5	2.25
5.6	5.5	0.1	0.01
5.5	5.5	0.0	0.00
4.1	5.5	−1.4	1.96
5.5	5.5	0.0	0.00
4.2	5.5	−1.3	1.69
5.6	5.5	0.1	0.01
5.4	5.5	−0.1	0.01
5.5	5.5	0.0	0.00
6.6	5.5	1.1	1.21
		$\Sigma (x - \bar{x})^2 =$	7.14

$$s = \sqrt{\frac{7.14}{10}} = .845$$

Thus the egg weight for frog 1 is expressed as 5.5 ± 0.845 mg.

Concerning the meaning of the standard evaluation, Fig. 16.2 has been drawn to show a perfect distribution curve. On such a
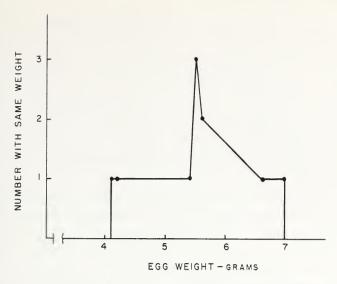
Fig. 16.1.   Frequency Distribution of Egg Weight—Insufficient Data.
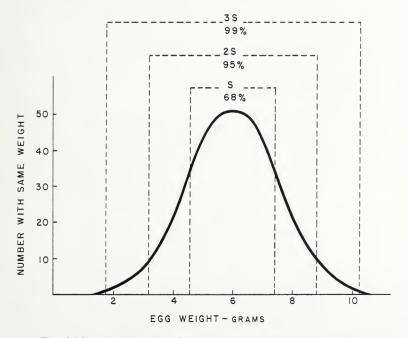
Fig. 16.2.   Frequency Distribution of Egg Weight—Sufficient Data.

curve the area between minus one standard deviation and plus one standard deviation includes 68 percent of all the measurements. Between $-2s$ and $+2s$ there are 95 percent, and between plus-and-minus three standard deviations 99 percent of the determinations are found. The standard deviation, then, affords a good indication of the variation of the measurements around the mean.

It should be noted that when the number of measurements is small, that is, below 30, $n - 1$ instead of n is used in the equation. So that to be absolutely correct in the frog-egg example the equation $s = \sqrt{\dfrac{(x - \bar{x})^2}{n - 1}}$ should be used which would give a standard deviation of .89. It can be appreciated that for numbers above 30, the difference between n and $n - 1$ is insignificant.

Unfortunately everyone does not use the same symbols. To avoid confusion the student should understand that standard deviation is most often represented, as has been done above, by the symbol s. The Greek letter σ is often used to stand for standard deviation when the number is more than 30, and the letter s used when the number is less than 30.

In the frog-egg problem, the experimental design called for the weighing of 8-10 eggs from each animal. The eggs from 12 frogs were weighed and the mean value and standard deviation for each animal calculated. The results were as follows:

Frog	Eggs	$\bar{x} \pm s$
1	10	$5.5 \pm .89$
2	10	$5.8 \pm .76$
3	9	$5.4 \pm .64$
4	10	$5.6 \pm .81$
5	8	$5.3 \pm .70$
6	9	$5.6 \pm .36$
7	10	$5.9 \pm .69$
8	9	$5.6 \pm .34$
9	9	$5.4 \pm .46$
10	10	$5.7 \pm .77$
11	8	$5.6 \pm .61$
12	10	$5.8 \pm .55$
		Avg. 5.6

It is immediately apparent that the same mean is not obtained for each frog. The question then arises as to whether or not there is a **significant** difference between these means. To answer that question a few additional concepts must be introduced.

Standard Error of the Mean

It has been seen that the values for egg weights in each frog varied around the mean. In a similar manner the mean values of 12 frogs vary around the mean of these means. The standard error of the mean is a measure of the variation of those means. It is a measure of precision. The symbol for the standard error of the mean most commonly used is $S.E._{\bar{x}}$ although $S.E.M.$, or just $S.E.$, is also quite correct. It is calculated by the equation:

$$ S.E._{\bar{x}} = \frac{s}{\sqrt{n-1}} $$

Again $n-1$ is used because the number of values, i.e., means, is less than 30. Using the above equation and substituting the data for the 12 frogs a standard error of the mean is found to be .055. Thus it can be said that the average egg weight for that particular species of frog is $5.6 \pm .055$.

Student's t Test

A man named Gossett, writing under a pseudonym of Student, devised a test of significance for small samples, that is, less than 30. This has become known as Student's t test which is widely used to evaluate the significance of the difference between two means. The following equation is used:

$$ t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{(S.E._{\bar{x}_1})^2 + (S.E._{\bar{x}_2})^2}} $$

Student's t Distribution

After the value of t has been determined by use of the above equation, Table 16.1 is used to determine the probability, P, that the difference between the means occurred by chance alone. Generally, a P value of .02, or in some cases .05, is taken as the dividing point

between significance and lack of significance. In other words, if cal-
culations disclose a P value of .02 there are only 2 chances in 100
that the difference between the means of samples of a certain size
occurred by chance alone. Or to put it in the reverse form, there
are 98 chances out of a 100 that some cause other than chance is re-
sponsible for the difference. This is said to be significant. Clearly,
the lower the P value the greater the significance.

It was noted that the mean weight for the eggs of each frog varied
in the group of 12 (cf. tabulation on page 300). Using Student's t test
and distribution we can now determine whether these differences
are statistically significant. The lowest mean weight was recorded
for frog 5 at 5.3 mg with a standard deviation of .70. Frog 7 had the
heaviest eggs: $\bar{x} = 5.9$ mg, $s = .69$. To determine t, as indicated
by the equation, it is necessary to calculate the standard error of the
mean for the two frogs, thus:

$$S.E._{\bar{x}} = \frac{s}{\sqrt{n-1}} = \frac{.70}{7} = .267 \text{ (frog 5)}$$

$$= \frac{.69}{9} = .33 \text{ (frog 7)}$$

$$\text{Then,} \quad t = \frac{5.9 - 5.3}{\sqrt{.33^2 + .265^2}} = 1.41$$

Table 16.1 is now used to determine the **probability, P.** In the table,
n stands for **degrees of freedom,** that is the sum of $n - 1$ for the two
samples being compared. In this case $n = 9 + 7 = 16$. Examination
of Table 16.1 discloses that for $n = 16$, and $t = 1.41$, P lies between
.2 and .1. It was stated above that to be considered significant P
must be less than .02. Clearly, then, the mean values of 5.9 and 5.3
under these conditions cannot be said to differ significantly one from
the other.

It was to be expected that eggs from frogs of the same size and of
the same species would have eggs of similar weight. This was con-
firmed by the data. But there is still the question as to whether or
not the eggs from this particular species are different from those of
another species. The mean for 12 frogs was found to be 5.6 mg with
a standard error of the mean of .055.

TABLE 16.1. Student's *t* Distribution *

n	Probability												
	.9	.8	.7	.6	.5	.4	.3	.2	.1	.05	.02	.01	.001
1	.158	.325	.510	.727	1.000	1.376	1.963	3.078	6.314	12.706	31.821	63.657	636.619
2	.142	.289	.445	.617	.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925	31.598
3	.137	.277	.424	.584	.765	.978	1.250	1.638	2.353	3.182	4.541	5.841	12.941
4	.134	.271	.414	.569	.741	.941	1.190	1.533	2.132	2.776	3.747	4.604	8.610
5	.132	.267	.408	.559	.727	.920	1.156	1.476	2.015	2.571	3.365	4.032	6.859
6	.131	.265	.404	.553	.718	.906	1.134	1.440	1.943	2.447	3.143	3.707	5.959
7	.130	.263	.402	.549	.711	.896	1.119	1.415	1.895	2.365	2.998	3.499	5.405
8	.130	.262	.399	.546	.706	.889	1.108	1.397	1.860	2.306	2.896	3.355	5.041
9	.129	.261	.398	.543	.703	.883	1.100	1.383	1.833	2.262	2.821	3.250	4.781
10	.129	.260	.397	.542	.700	.879	1.093	1.372	1.812	2.228	2.764	3.169	4.587
11	.129	.260	.396	.540	.697	.876	1.088	1.363	1.796	2.201	2.718	3.106	4.437
12	.128	.259	.395	.539	.695	.873	1.083	1.356	1.782	2.179	2.681	3.055	4.318
13	.128	.259	.394	.538	.694	.870	1.079	1.350	1.771	2.160	2.650	3.012	4.221
14	.128	.258	.393	.537	.692	.868	1.076	1.345	1.761	2.145	2.624	2.977	4.140
15	.128	.258	.393	.536	.691	.866	1.074	1.341	1.753	2.131	2.602	2.947	4.073
16	.128	.258	.392	.535	.690	.865	1.071	1.337	1.746	2.120	2.583	2.921	4.015
17	.128	.257	.392	.534	.689	.863	1.069	1.333	1.740	2.110	2.567	2.898	3.965
18	.127	.257	.392	.534	.688	.862	1.067	1.330	1.734	2.101	2.552	2.878	3.922
19	.127	.257	.391	.533	.688	.861	1.066	1.328	1.729	2.093	2.539	2.861	3.883
20	.127	.257	.391	.533	.687	.860	1.064	1.325	1.725	2.086	2.528	2.845	3.850
40	.126	.255	.388	.529	.681	.851	1.050	1.303	1.684	2.021	2.423	2.704	3.551
120	.126	.254	.386	.526	.677	.845	1.041	1.289	1.658	1.980	2.358	2.617	3.373
∞	.126	.253	.385	.524	.674	.842	1.036	1.282	1.645	1.960	2.326	2.576	3.291

Degrees of Freedom

Assume that another species shows a mean of 5.0, $S.E._{\bar{x}} = .120$ and $n = 6$.

$$t = \frac{5.6 - 5.0}{\sqrt{.055^2 + .120^2}} = 4.41$$

In this case the degrees of freedom are 16, that is, $(11 + 5)$. P is less than .001. Beyond question there is a highly significant difference between these two means, and therefore with great confidence we can say that there is a different egg weight in the two species of frogs. Even though only 6 frogs of the second species were used, the statistical analysis of the data makes it clear that there would be no point in using any more.

Student's t distribution table, then, not only is of value in determining whether or not there is a significant difference between two mean values, but it also provides an indication of how many animals must be included in each group. The above discussion should have made it clear that if the variation within each group is small and if the difference between the two means is large a relatively small number of animals suffices. Conversely, with large variation and a small difference between the means many more determinations must be made. If from prior knowledge, or preliminary studies, one has a fair idea of his mean values, the standard deviation, and the approximate difference between the means of the groups being investigated, then it is possible to determine how many measurements must be included in each group in order to obtain a certain level of significance. Tables have been prepared for this purpose.

Chi-Square Test

The chi-square test is used when the data fits into the yes or no type of categories. This test answers the question as to whether something did or did not have an effect. The classic example of coin tossing is always used to illustrate the chi-square test.

If a coin is perfectly balanced or true, it is to be expected that with repeated tossings it will land heads 50% of the time and tails the other 50%. The chi-square test is based on the **null hypothesis** which means that if there is no imbalance, if there is truly no difference between the two sides, then each side will come up the same number of times.

The coin is tossed many times, for convenience say 50. It is antici-
pated, according to the null hypothesis, that the coin will land heads
25 times and tails the same number. Then:

$$x^2 = \Sigma \frac{(\text{Observed} - \text{Expected})^2}{\text{Expected}}$$

The coin actually landed 32 times heads and only 18 times tails,
therefore:

$$
\begin{array}{ccc}
(\text{Heads}) & & (\text{Tails}) \\
x^2 = \dfrac{(32 - 25)^2}{25} & + & \dfrac{(18 - 25)^2}{25} \\
= 1.96 & + & 1.96 \quad = 3.92
\end{array}
$$

Table 16.2 shows the chi-square distribution. In this table, it
must be noted, n equals the degrees of freedom as in Table 16.1,
but here it represents not the number of measurements, but rather
the independent values. In the above equation it is obvious that
there is only one independent value. If it was only known that heads
had come up 32 times, then the number of tails could have been
calculated. In short, in this illustration there is but one degree of
freedom. In Table 16.2 a value of 3.92 with one degree of freedom
shows a probability of only slightly less than .05. As discussed previ-
ously, this is barely significant. It may mean that the coin was not
tossed a sufficiently great number of times. If a coin is tossed another
50 times and the same ratio obtained, there would now be 64 heads
and 34 tails. With these figures a P of less than .01 results. This is
highly significant.

A typical situation in which the chi-square test would be used is
the evaluation of a drug. Assume that the drug under investigation
is thought to combat fever. The drug is given to one group of patients
with fever; the control group receives an impotent substitute, a so-
called **placebo.** The raw data at the end of the observation is:

NUMBER OF PATIENTS

	Fever	Normal	Total	Percent Normal
Control	46	7	53	13.2
Drug	12	31	43	72.2
Total	58	38	96	39.6

TABLE 16.2. Chi-Square Distribution *

n	Probability													
	.99	.98	.95	.90	.80	.70	.50	.30	.20	.10	.05	.02	.01	.001
1	.03157	.03628	.00393	.0158	.0642	.148	.455	1.074	1.642	2.706	3.841	5.412	6.635	10.827
2	.0201	.0404	.103	.211	.446	.713	1.386	2.408	3.219	4.605	5.991	7.824	9.210	13.815
3	.115	.185	.352	.584	1.005	1.424	2.366	3.665	4.642	6.251	7.815	9.837	11.345	16.268
4	.297	.429	.711	1.064	1.649	2.195	3.357	4.878	5.989	7.779	9.488	11.668	13.277	18.465
5	.554	.752	1.145	1.610	2.343	3.000	4.351	6.064	7.289	9.236	11.070	13.388	15.086	20.517
6	.872	1.134	1.635	2.204	3.070	3.828	5.348	7.231	8.558	10.645	12.592	15.033	16.812	22.457
7	1.239	1.564	2.167	2.833	3.822	4.671	6.346	8.383	9.803	12.017	14.067	16.622	18.475	24.322
8	1.646	2.032	2.733	3.490	4.591	5.527	7.344	9.524	11.030	13.362	15.507	18.168	20.090	26.125
9	2.088	2.532	3.325	4.168	5.380	6.393	8.343	10.656	12.242	14.684	16.919	19.697	21.666	27.877
10	2.558	3.059	3.940	4.865	6.179	7.267	9.342	11.781	13.442	15.987	18.307	21.161	23.209	29.588
11	3.053	3.609	4.575	5.578	6.989	8.148	10.341	12.899	14.631	17.275	19.675	22.618	24.725	31.264
12	3.571	4.178	5.226	6.304	7.807	9.034	11.340	14.011	15.812	18.549	21.026	24.054	26.217	32.909
13	4.107	4.765	5.892	7.042	8.634	9.926	12.340	15.119	16.985	19.812	22.362	25.472	27.688	34.528
14	4.660	5.368	6.571	7.790	9.467	10.821	13.339	16.222	18.151	21.064	23.685	26.873	29.141	36.123
15	5.229	5.985	7.261	8.547	10.307	11.721	14.339	17.322	19.311	22.307	24.996	28.259	30.578	37.697
16	5.812	6.614	7.962	9.312	11.152	12.624	15.338	18.418	20.465	23.512	26.296	29.633	32.000	39.252
17	6.408	7.255	8.672	10.085	12.002	13.531	16.338	19.511	21.615	24.769	27.587	30.995	33.409	40.790
18	7.015	7.906	9.390	10.865	12.857	14.440	17.338	20.601	22.760	25.989	28.869	32.346	34.805	42.312
19	7.633	8.567	10.117	11.651	13.716	15.352	18.338	21.689	23.900	27.204	30.144	33.687	36.191	43.820
20	8.260	9.237	10.851	12.443	14.578	16.266	19.337	22.775	25.038	28.412	31.410	35.020	37.566	45.315
21	8.897	9.915	11.591	13.210	15.445	17.182	20.337	23.858	26.171	29.615	32.671	36.343	38.932	46.797
22	9.512	10.600	12.338	14.041	16.314	18.101	21.337	24.939	27.301	30.813	33.924	37.659	40.289	48.268
23	10.196	11.293	13.091	14.848	17.187	19.021	22.337	26.018	28.429	32.007	35.172	38.968	41.638	49.728
24	10.856	11.992	13.848	15.659	18.062	19.943	23.337	27.096	29.553	33.196	36.415	40.270	42.980	51.179
25	11.524	12.697	14.611	16.473	18.940	20.867	24.337	28.172	30.675	34.382	37.652	41.566	44.314	52.620
26	12.198	13.409	15.379	17.292	19.820	21.792	25.336	29.246	31.795	35.563	38.885	42.856	45.642	54.052
27	12.879	14.125	16.151	18.114	20.703	22.719	26.336	30.319	32.912	36.741	40.113	44.140	46.963	55.476
28	13.565	14.847	16.928	18.939	21.588	23.647	27.336	31.391	34.027	37.916	41.337	45.419	48.278	56.893
29	14.256	15.574	17.708	19.768	22.475	24.577	28.336	32.461	35.139	39.087	42.557	46.693	49.588	58.302
30	14.953	16.306	18.493	20.599	23.364	25.508	29.336	33.530	36.250	40.256	43.773	47.962	50.892	59.703

Degree of Freedom

* Table 16.2 is abridged from Table IV for Fisher and Yates, "Statistical Tables for Biological, Agricultural and Medical Research," published by Oliver & Boyd Ltd, Edinburgh. By permission of the authors and publishers.

It can be seen that for the totals of the two groups 39.6 percent ($38/96 \times 100$) of all the patients had no fever at the end of the treatment period. If no difference (null hypothesis) exists between the two groups then it would be expected that in the control group 21 patients would be normal ($53 \times 39.6\%$) and 32 patients ($53 - 21$) would still have fever. The same procedure is carried out for the drug group and the data may then be arranged as follows:

NUMBER OF PATIENTS

	Fever		Normal		
	Obs.	Exp.	Obs.	Exp.	Total
Control	46	32	7	21	53
Drug	12	26	31	17	43
Total	58	58	38	38	96

$$x^2 = \frac{(46-32)^2}{32} + \frac{(12-26)^2}{26} + \frac{(7-21)^2}{21} + \frac{(31-17)^2}{17}$$
$$= 6.32 \qquad 7.55 \qquad 9.34 \qquad 11.52$$
$$= 34.73$$

There is still but one degree of freedom because after only one value was calculated the other three could be obtained by subtraction. Using Table 16.2 it is found that for one degree of freedom and chi-square values of 34.73, P is less than .001. This means that there is a probability of less than 1 in a 1000 that such a large difference between the two groups occurred by chance. In short, there is a highly significant difference between the two groups, and it may be concluded that the drug under investigation lowers fever.

The chi-square test may be used to handle data in many other situations than these discussed above. For an excellent discussion of this subject the reader is referred to "Introduction to Biostatistics" by Huldah Bancroft.

Correlation

It will be recalled that on the data sheet for the frog-egg example there was a space to record the weight of the frog. Twelve frogs were used and the mean weight of the eggs from each ascertained. The question now arises as to whether or not there is any relation-

ship, that is correlation, between the weight of the frog and the weight of the eggs. The data is as follows:

Frog	Egg Weight (x) mg.	Frog Weight (y) gm.
1	5.5	125
2	5.8	140
3	5.4	123
4	5.6	127
5	5.3	121
6	5.6	130
7	5.9	141
8	5.6	128
9	5.4	124
10	5.7	138
11	5.6	217
12	5.8	136
	$\bar{x} = 5.6$	$\bar{y} = 130$

A rough idea as to whether or not there is a possible correlation is had by simply examining the data. When this is done, it is noted that the heavier frogs do indeed seem to have heavier eggs. Another way of examining such data is to plot it on a graph as has been done in Fig. 16.3. A very definite pattern is observed. All of the points do not fit exactly on the line and therefore it may be of importance to know how well they do fit. To do this the **correlation coefficient** is calculated.

$$\text{Correlation coefficient } r = \frac{\Sigma(x - \bar{x})(y - \bar{y})}{n \cdot s_x \cdot s_y}$$

To solve this equation it is necessary to subtract the means from the actual values for both sets of data. For Frog 1 it would be $5.5 - 5.6$ for the egg weight (x) and $125 - 130$ for the frog weight (y). The remainders would be -0.1 and -5. According to the equation these values are multiplied, and the product is $+0.5$. The same procedure is carried out for each frog and then all the products added together and divided by n times the standard deviation for the egg weight times the standard deviation for the frog weight. This works out as follows:

$$r = \frac{13.1}{12 \times 0.181 \times 6.95}$$
$$= \quad + .87$$

The range of possible r values is from -1 to $+1$. A value of $+1$ indicates a perfect positive correlation; -1 indicates a perfect negative correlation. In both cases all the points would lie exactly on a straight line when the data is plotted. In the frog example a value of $+.87$ resulted. This indicates that there is a high, but not perfect positive correlation between frog size and egg size, or to put it another way, the two values, egg weight and frog weight, change in the same direction.
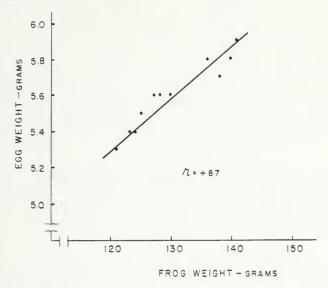
Fig. 16.3.   Correlation of Frog Weight and Egg Weight.

If a value of 0 is obtained, it means that there is absolutely no correlation between the two factors under study. That is to say, they are perfectly independent of one another. Where, then, is the dividing line between no correlation and a real correlation? There is no sharp line; it is rather one of degree. The closer the r value to 1 the higher the correlation.

SUMMARY

In any determination, or **measurement**, there may be two types of **error**, systematic and random. A systematic error is caused by a defect in the measuring system. **Random** error simply represents variation from the true value caused by factors which average out in a

large number of determinations. The experimental design minimizes systematic error by employing appropriate measuring instrumentation. Repeated determinations reduce the random error to within acceptable limits.

The raw data is subjected to biostatistical analysis. The **mean** is the arithmetic average. Variation about that mean is indicated roughly by the **range** and more precisely by the **standard deviation.** On the normal distribution curve 68 percent of the values are contained between plus and minus one standard deviation. Just as individual values vary, so do mean values of several groups. The **standard error of the mean** is a measure of the variation of those means.

Student's *t* test is used to evaluate the significance of the difference between two means. The value of *t* in relation to the number of measurements discloses the **probability** of such a difference in the means occurring purely by chance. The **chi-square** test of significance is used when the data is of the yes or no, change or no change type.

To determine whether or not there is a relationship between two factors the **correlation** is determined. A value of $+1$ indicates a perfect positive correlation, -1 a perfect negative correlation, and 0 no correlation or complete independence of the factors.

Problems

1. The plasma of a group of animals was analyzed and found to contain the following levels of potassium:

Animal no.	Plasma K mEq/l
1	4.2
2	5.0
3	4.7
4	4.9
5	4.4
6	4.6
7	4.5
8	4.7
9	4.3

a) Determine the mean and the range.
b) Determine the standard deviation and the standard error of the mean.

2. A second group of animals was treated with adrenal steroids and found to have the following levels of potassium:

Animal no.	Plasma K mEq/l
1	3.8
2	4.3
3	4.4
4	3.9
5	4.0
6	4.4
7	4.0
8	3.9

Is there a significant difference between these two groups of animals?

3. Describe an experiment in which the resulting data would be subjected to the chi-square test.

4. Calculate the correlation between the height and weight of the following individuals:

Individual no.	Height cm	Weight kg
1	175	75
2	168	75
3	170	84
4	176	90
5	166	70
6	175	80
7	160	62
8	178	92
9	172	74
10	169	71
11	164	66
12	180	92

a) Is there a correlation?
b) Is it negative or positive?
c) What is the relative degree of correlation?

Additional Reading

Bancroft, H., "Introduction to Biostatistics," Hoeber-Harper, New York, N. Y., 1957.

Croxton, F. E., "Elementary Statistics," Dover Publications, Inc., New York, N. Y., 1953.

Dixon, W. J., and Massey, F. J., Jr., "Introduction to Statistical Analysis," 2d ed., McGraw-Hill Book Co., New York, N. Y., 1957.

Fisher, R. A., "The Design of Experiments," Oliver & Boyd, Ltd., Edinburgh, Scotland, 1949.

Moore, F. J., Cramer, F. B., and Knowles, R. G., "Statistics for Medical Students," Blakiston Company, New York, N. Y., 1951.

Snedecor, G. W., "Statistical Methods," College Press, Iowa State, Ames, Iowa, 1950.

CHAPTER 17

MOLECULAR ACTIVITY

A KNOWLEDGE of molecular activity is absolutely fundamental to an understanding of cell physiology. Broadly speaking, a molecule is any minute particle. More specifically, it is the unit of matter, a unit capable of independent existence. Some molecules are composed of two atoms. These atoms may be identical, as in the case of N_2, or they may be different as in NaCl. There are also molecules composed of but one atom, for example Hg.

Matter, which as has been noted is made up of molecules, exists in one of three states: gas, liquid, or solid. In the **gaseous state** the molecules are not only in constant motion, but they move very rapidly and in a purely random fashion. The molecules exert so little influence upon one another that a mass of gas has no shape or volume of its own. It simply fills any container and thus assumes the shape and volume of that container. In contrast, in the **liquid state** the molecules still move but do so within the boundaries of interacting forces; thus a liquid has no definite shape, but it does have a very definite volume. A liquid placed in a container will not necessarily fill it as will a gas, but it will take the shape of the part of the container it does fill. Finally, in the **solid state** the molecules attract each other with such force that they remain fixed in a set pattern. A solid, therefore, has both a definite shape and volume.

THE GAS LAWS

Avogadro's Law

In 1811 an Italian professor of mathematical physics published a hypothesis which has been confirmed and has stood the test of time

so that it is now known as a law. Avogadro stated that equal volumes of different gases under the same temperature and pressure contain the same number of molecules. To put it another way, 1 gram-molecular weight of a gas at a given temperature and pressure occupies the same volume as 1 gram-molecular weight of any other gas under the same condition. A gram-molecular weight is, of course, the molecular weight of the substance expressed in grams. $0°C$ and 1 atmosphere, or 760 mm Hg, constitute what is known as **standard conditions**, that is, standard temperature and pressure and is abbreviated **STP.** It has been found that at STP, 1 gram-molecular weight of any gas occupies a volume of 22.4 liters. One gram-molecular weight of any substance also contains the same number of molecules. This number, known as **Avogadro's number**, is 6.02×10^{23}. If these facts are combined, it is found that 1 gram-molecular weight of any gas contains 6.02×10^{23} molecules and, under standard conditions, occupies a volume of 22.4 liters. Therefore:

$$V \propto n \text{ (moles)}$$

Boyle's Law

In a gas the molecules are normally far apart, so that there remains much empty space. It is for this reason that gases are so easily compressible. This fact was quite apparent to the English chemist and physicist, Robert Boyle, as long ago as 1662. He found experimentally that there is a definite relationship between the pressure under which a gas is subjected and its volume. Thus he noted that if the pressure is doubled, the volume is halved. Boyle's law, then, simply states that the volume of a gas varies inversely with the pressure. Therefore:

$$V \propto \frac{1}{P} \text{ and } \frac{V_1}{V_2} = \frac{P_2}{P_1}$$

Charles' Law

In 1787 J. A. C. Charles, a French physicist, found that all gases expand equally when heated, if the pressure is kept constant. Therefore:

$$V \propto T \text{ and } \frac{V_1}{V_2} = \frac{T_1}{T_2}$$

It must be understood that T in the above equation refers to the **absolute temperature.** Absolute temperature equals the centigrade temperature plus approximately 273°C. This figure comes from the so-called absolute zero which is theoretically the lowest temperature possible. It has been calculated to be −273.16°C. It has been seen that at STP, which is 0°C and 760 mm Hg, the gram-molecular weight of a gas will occupy a volume of 22.4 liters. To double this volume to 44.8 liters it is necessary to raise the temperature to 273.16°C. Theoretically if the temperature were lowered to absolute zero (−273.16°C), the gas would have no volume.

Combined Gas Laws

The gas laws may be combined into one general equation known as the **general ideal gas equation.** It is:

$$V \propto \frac{nT}{P} \text{ or } \frac{PV}{nT} = R \text{ (a constant)}$$

$$\text{therefore, } PV = nRT$$

Dalton's Law

John Dalton, an English physicist and chemist, made many contributions to his field of study. Among them was the theory he published in 1802 which states that the pressure of a mixture of gases is equal to the sum of the pressures of the gases in the mixture. In other words, each gas simply exerts part of the pressure of the whole. Accordingly, this relationship has become known as Dalton's law of **partial pressures.** It is found that each gas in a mixture exerts the same pressure as it would were it to occupy the same volume alone. Thus:

$$\text{The total pressure } P = p_1 + p_2 + p_3 \cdots \cdots p_n$$

Henry's Law

William Henry was another English physicist who studied the behavior of gases. At the beginning of the last century he showed that the volume of a gas absorbed by a liquid is directly proportional to the pressure of the gas. It follows that if the gas exists in a mixture of gases the amount (weight) of that particular gas absorbed in a liquid will be proportional to its partial pressure. Henry's law, then, states that the solubility of each gas in a mixture of gases is propor-

tional to the partial pressure of each gas at equilibrium. It should be noted that the volume of gas taken up by unit volume of liquid is the same under all pressures; it is the weight that changes. With greater pressure there is simply more gas in the same volume. It should also be emphasized that Henry's law is only applicable if there is no chemical reaction between the gas and the liquid. It has reference only to a physical solution.

Graham's Law

In 1832 Thomas Graham, an English chemist, stated that the relative speeds of diffusion of gases are inversely proportional to the square roots of their relative densities. The density of a gas is indicated by its molecular weight. Thus, since the molecular weight of oxygen is 16 times greater than that of hydrogen, hydrogen diffuses four times faster than does oxygen. Perhaps some thought should be given to the word **diffusion.** To diffuse means to spread, to flow out, to extend in all directions, to disperse. Thus, when two gases are placed together, it is found that they spread and extend and disperse in all directions. The same applies to liquids, but since the molecular activity in a liquid is less than in a gas the rate of diffusion is lower.

THE LIQUID STATE

In the introduction to this chapter it was mentioned that in gases and liquids the molecules are in constant and random motion. The average velocity for the molecules of any particular gas or liquid varies with the temperature. The term "average velocity" is used because at any temperature most of the molecules have the same velocity but a few have higher and a few lower velocities. As the temperature is raised the percentage of molecules having higher velocities increases and thus so does the average velocity. A knowledge of this relationship between temperature and molecular activity is essential for an understanding of vapor pressure.

Vapor Pressure

Figure 17.1 shows a container partially filled with a liquid. Two molecules are shown; one at the surface and the other deep in the

liquid. It must be kept in mind that there is a large number of other molecules dispersed throughout the liquid, all exerting forces of attraction on each other and on the two molecules shown. A difference in the resultant force acting on each molecule is immediately apparent. The molecule deep in the liquid is surrounded by other molecules equally on all sides and thus the attractive forces balance out and the resultant is zero. But the molecule at the surface is only partially surrounded and therefore the resultant, as shown by the arrow, is downward. This downward pull is opposed by the inherent motion of the molecule. If the velocity of that particular molecule is average, or less than average, the resulting attraction of the other molecules will keep it from flying out of the liquid. On the other hand, if this particular molecule happens to be one with a higher than average velocity, it may very well be propelled out of the liquid despite the attractive forces of the surrounding molecules. If the container is an open one, the escaped molecule may be lost to the surrounding space. There remains less liquid. This conversion of liquid to a gas is termed **evaporation**.

Fig. 17.1. Molecular Attraction. In the submerged molecule forces resulting from attraction by other molecules are about equal. At the surface the resultant of the molecular forces holds the molecule in the solution.

Since the percentage of high velocity molecules increases with the temperature, the number of molecules that will escape from the surface of a liquid is also a function of the temperature. This simply is another way of expressing the well known fact that the higher the temperature the faster a liquid will evaporate.

If the liquid is in a closed container, the high velocity molecules will still fly out of the liquid, but they will then ricochet off the sides and top of the container. They may be reflected back into the liquid or they may stay in the gas above the liquid. If the temperature is maintained constant, an equilibrium will soon be reached in which the number of molecules leaving the liquid is balanced by the number returning. If a manometer is attached to the container as shown

in Fig. 17.2, the pressure of the gas, that is the vapor, above the liquid may be read directly. This is the **vapor pressure.** Clearly, if the temperature is raised, more molecules will escape into the vapor until a new equilibrium is reached. At this point the vapor pressure will be found to be higher. It may thus be concluded that the vapor pressure of any liquid varies directly with the temperature. It is not, however, a linear relationship. As shown in Fig. 17.3, the higher the temperature the more rapid is the increase in vapor pressure.
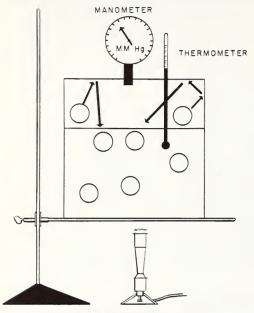
Fig. 17.2. Method of Determining Vapor Pressure.

Even if a container is open, there is still pressure exerted on the surface of the liquid. This is the atmospheric pressure; it is equal to 760 mm Hg at sea level. Fig. 17.3 shows that when water is heated at 100°C the vapor pressure is 760 mm Hg. Now the water molecules freely escape and the liquid is said to **boil.** Accordingly, water may be made to boil at any temperature by varying the overlying pressure. For example, if the pressure is reduced to 17.5 mm Hg water will boil at 20°C. This explains why it is difficult to cook at extremely high elevations. It also provides an explanation of the function of a pressure cooker.

Surface Tension

If one carefully observes a drop as it forms at the end of a pipette he will note that it clings to the glass. Then he will see the drop grow larger, become elongated and finally break away from the tip. As soon as the drop begins to fall free, it forms a typical spherical shape. Several principles are illustrated by this simple demonstra-
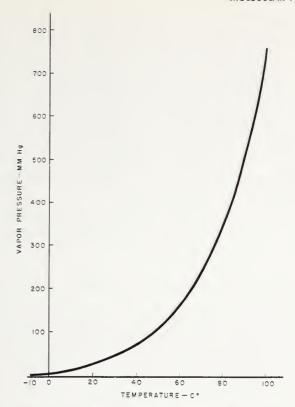
Fig. 17.3. Relationship of Temperature and Vapor Pressure.

tion. First, it is obvious that there is a clinging of the drop to the pipette. It adheres. This shows an attraction between the water molecules and those of the glass pipette. Such attraction between dissimilar molecules is termed **adhesion**.

It is also observed that as the drop increases in size the force of gravity causes it to become elongated. It stretches but does not break. Evidently, within the drop, forces must exist which hold the drop together. As has already been emphasized in the discussion of vapor pressure, molecules exert attractive forces. Such attraction between similar molecules is termed **cohesion**.

If reference is made again to Fig. 17.1 it will be recalled that when a molecule is at the surface of a liquid the resultant force of cohesive attraction tends to pull that molecule away from the surface. If one

visualizes this pattern existing everywhere on the surface, then it will become clear that a drop of liquid is pulled together into a shape with the smallest possible surface area. This is a sphere. These inward forces on the surface molecules oppose stretching or breaking, and constitute what has been termed **surface tension.** It is measured by determining the force necessary to stretch or break that surface.

The surface tension is a characteristic of each liquid. For example, the surface tension of water, expressed in dynes per centimeter, is about 73. However, for ethyl alcohol it is but 22. These figures simply reflect the relative cohesive forces of the different types of molecules which make up the liquids.

In discussing vapor pressure the influence of temperature was emphasized. Since higher temperatures increase the average velocity of the molecules, more of them escape from the surface. This is simply to say that heat produces forces which oppose cohesive attraction. Accordingly, the surface tension of any liquid is inversely related to the temperature.

The role of surface tension in basic physiological mechanisms is becoming ever more apparent. Enzyme action, the movement of liquids through cell membranes, and many other phenomena are definitely influenced by the **interfacial** tension which exists between any two liquids.

SOLUTIONS

Thus far, matter has been considered in the pure states, solid, liquid, and gas. A solution is a mixture of two or more substances. It may be a solution of any combination of the three states of matter. The solution most often dealt with in physiological work is that of a solid dissolved in a liquid. The liquid, that is the dissolving agent, is called the **solvent;** the substance dissolved in the solvent is termed the **solute.**

Concentration

In the scientific literature certain expressions are commonly used to express the concentration of solutions. Perhaps most often concentration is simply stated as weight of solute per unit volume of solvent. For example, substances dissolved in the blood are some-

times expressed as milligrams percent, that is, the number of milli-
grams of the particular substance existing in 100 milliliters of blood.
Concentration may also be expressed in terms of **molarity**. A 1 molar
solution contains 1 gram-molecular weight of any substance in 1 liter
of total solution. **Normality** expresses concentration in terms of gram
equivalent weights divided by the valence. Thus a 1 normal solu-
tion of $CaSO_4$ contains but $\frac{1}{2}$ gram-molecular weight per liter
whereas a 1 normal solution of $NaCl$ contains 1 gram-molecular
weight per liter. The expression "equivalents," or "milli-equiva-
lents" is often used. Thus one may see the concentration of potas-
sium in plasma stated as 4 mEq/l, or the same concentration could
be expressed as 15.6 mg percent $\left[\frac{(4 \times 39)}{10} \right]$

Osmosis

If two solutions of different concentration are separated by a
semipermeable membrane, the solvent from the less concentrated will
move through the membrane into the more concentrated solution.
This process is termed osmosis. Actually it is not necessary to have
a membrane present. For example, pure water may be carefully
stratified over a salt solution. Here again, the water molecules will
pass into the concentrated salt solution. But to measure the pressure
exerted by the osmotic processes, that is, the **osmotic pressure**, a
semipermeable membrane is essential. Such a membrane is one
through which the solvent, but not the solute, can pass.

Figure 17.4 shows a simple arrangement in which it is possible to
measure osmotic pressure. The bell at the end of the tube is closed
by a semipermeable membrane. The bell contains a water solution,
the solute of which cannot pass the membrane. The beaker contains
pure water. Water molecules are freely diffusable through the mem-
brane. In this system the solution will ascend in the tube to a cer-
tain height and then stop. The height of this column of fluid is
an approximation of the osmotic pressure. At equilibrium the hydro-
static pressure of the column balances the force of the water moving
through the membrane. Thus, if the column is measured it will ap-
proximate the osmotic pressure expressed in centimeters H_2O. The
term approximate is used because as the water enters the solution in
the bell it dilutes it and thus lowers its osmotic pressure. Therefore,

the osmotic pressure at equilibrium is not the same as before the
dilution began. A true measure of the osmotic pressure can be had
by raising the pressure over the concentrated solution so as to pre-
vent elevation in the tube. The pressure which just prevents water
movement is the true osmotic pressure of the solution. The osmotic
pressure of a solution is proportional to its molecular concentration.
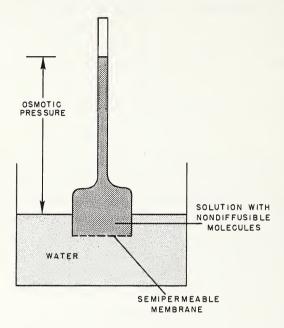
Fig. 17.4. A Method for Estimating Osmotic Pres-
sure. The fluid will rise in the tube until the hydro-
static pressure equals the osmotic pressure.

Most membranes are not completely impermeable to one molecule
while remaining permeable to another. Thus if two solutions are
separated by such a membrane, solvent will pass into the more con-
centrated solution, but solute will also pass in the opposite direction.
There will be, however, a difference in rate dependent upon the
relative permeability for the two types of molecules. Usually, the
solvent passes more rapidly. Thus at first there will be an increase
in volume of the more concentrated solution, but at equilibrium the
volume and concentration on both sides of the membrane will be
identical.

Ionization

The osmotic pressure of a solution depends upon the number of molecules present. In solutions of organic compounds, the osmotic pressure is proportional to the molecular weight. But when one compares a solution of salt with a solution of sugar, it is found that at equal gram-molecular concentrations the sodium chloride solution has twice the osmotic pressure as does that of glucose. The explanation lies in the fact that osmotic pressure depends upon the number of particles present in the solution and when a molecule of sodium chloride is dissolved two particles result. These particles are called **ions** and the process is known as ionization. Other molecules may produce more than two ions.

In Fig. 9.1 (page 165) the atomic structure of hydrogen, lithium, and oxygen are shown. It is seen that the number of protons in the nucleus of each atom is the same as the number of electrons surrounding each atom. These fundamental units of matter are charged; protons positive and electrons negative. Thus, when the atom contains the same number of each, the entire atom is neutral; it has no charge. Atoms interreact to form compounds; in the case of sodium and chlorine, sodium chloride results. When these atoms form a molecule of NaCl, there is a movement of one electron from the sodium atom to the chlorine atom. Henceforth each component of the compound is termed an ion and has a charge. The sodium ion now has one more proton than electrons and therefore, it has a positive charge and is written Na^+. The chloride ion has an extra electron and thus is negative. It is written Cl^-.

When organic molecules, such as glucose, are dissolved the entire molecule clings together in one particle. But when inorganic molecules, such as sodium chloride, are dissolved, the components, the ions, separate. This process of separation is termed **dissociation.** If an electrical current is applied to the salt solution, the charged particles will move to the electrodes and give up their charge. This process completes the circuit between the electrodes and thus the solution is said to conduct a current and therefore is called an electrolytic solution. The salt is the **electrolyte.** Since organic molecules do not dissociate into charged particles, such a solution will not conduct a current.

Acids and Bases

Hydrogen has but one proton and one electron; thus when it reacts, say with chlorine, it gives up its electron and the resulting hydrogen ion, H^+, is nothing more than a single proton. If this fact is kept in mind, then the current concept of acids and bases becomes understandable. It is now accepted that an **acid** is any substance capable of giving up a proton, that is, the hydrogen ion. A **base** is any substance capable of accepting a proton. A base, then, could be any one of a great number of substances, molecular or ionic. A few examples are, OH^-, Cl^-, NO_3^-, and NH_3. There are also many acids which have in common the availability of the hydrogen ion, that is, the proton. According to this concept, the strength of an acid or a base depends upon the facility with which it gives up or accepts the proton.

SUMMARY

Matter exists in one of three states: **gas, liquid,** or **solid.** In a gas the molecules move so freely that it has no independent shape or volume. The molecules of a liquid exert sufficient interattraction so that a liquid does have a definite volume. This molecular attraction in a solid is so great that both the shape and volume are fixed.

Avogadro's law states that equal volumes of different gases under the same temperature and pressure contain the same number of molecules. According to **Boyle's law** the volume of a gas varies inversely with the pressure. **Charles' law** states that the volume of a gas varies directly with the **absolute** temperature. In a mixture of gases each exerts the same pressure as it would were it to occupy the same volume alone. This is **Dalton's law of partial pressures. Henry's law** states that if there is no chemical reaction between a gas and a liquid, the amount of gas absorbed in the liquid will be proportional to its partial pressure. And according to **Graham's law** the relative speeds of diffusion of gases are inversely proportional to the square roots of their relative densities.

Molecules constantly escape from a liquid to form a gas. The pressure exerted by this gas at equilibrium is termed the **vapor pressure.** It varies directly with the temperature. The attraction between similar molecules, **cohesion,** in a liquid tends to prevent escape from the

surface. There is therefore **surface tension.** The surface tension varies from liquid to liquid. In any liquid, the surface tension is inversely related to the temperature.

A **solution** is a mixture of two or more substances. Concentration of the **solute** in the **solvent** may be expressed as weight of solute per unit of solvent, or in terms of **molarity** or **normality.** The process by which a liquid moves through a semipermeable membrane into a solution of higher concentration is termed **osmosis.** The **osmotic pressure** of a solution is proportional to its molecular concentration. Actually, it depends upon the number of particles present. Molecules of electrolytes dissociate into two or more **ions** in solution. Due to an imbalance between protons and electrons, ions have a charge and can therefore conduct a current. An **acid** is any substance, molecular or ionic, capable of giving up a proton. A **base** is one capable of accepting a proton.

Problems

1. If a gas occupies a volume of 1.5 liters at STP what will its volume be at a pressure of 755 mm Hg and a temperature of 20°C?
2. Air contains about 80 percent nitrogen and 20 percent oxygen. If the atmospheric pressure is 760 mm Hg what is the partial pressure of oxygen?
3. Which gas, oxygen or nitrogen, will diffuse faster? Why?
4. Explain why vapor pressure is influenced by temperature.
5. How can surface tension be measured?
6. How would you prepare a one normal solution of NaH_2PO_3?
7. Would the osmotic pressure be the same in a one molar solution of glucose as it is in a one molar solution of NaCl? Explain.

Additional Reading

Bull, H. B., "Physical Biochemistry," 2nd ed., John Wiley & Sons, Inc., New York, N. Y., 1951.

Crockford, H. D., and Knight, S. B., "Fundamentals of Physical Chemistry for Premedical Students," John Wiley & Sons, Inc., New York, N. Y., 1950.

Daniels, F., "Outlines of Physical Chemistry," John Wiley & Sons, Inc., New York, N. Y., 1948.

Hitchcock, D. I., "Physical Chemistry for Students of Biology and Medicine," Charles C Thomas, Springfield, Ill., 1932.

Johlin, J. M., "Introduction to Physical Biochemistry," 2nd ed., Paul B. Hoeber, Inc., New York, N. Y., 1949.

ELECTRICITY

THE ACTIVITY of muscle or a gland may be easily observed; one needs only to watch a muscle contract and a gland secrete to know that these tissues are active. But when nervous tissue is propagating an impulse, there is no such obvious indication. More subtle means of observation are required. The activity of nervous tissue is signaled by alterations in electrical potential. Accordingly, an understanding of the nature and use of electrical energy is of primary importance to the physiologist. As a matter of fact, every living cell exhibits a potential difference between its inside and outside. Further, electrical stimuli are most often the stimuli of choice in experimental work since they can be so easily and so precisely controlled. And finally, the present trend in physiological recording is to make use of electrical energy. For these reasons this chapter is concerned with a consideration of electricity.

CHARGED PARTICLES

The term **electricity** is difficult to define. Electricity is a fundamental quantity of nature. It consists of elementary particles termed **electrons** and **protons,** and is characterized by the fact that it gives rise to a field of force possessing potential energy. Likewise, when electricity moves in a stream, that is, a current, it gives rise to a magnetic field of force with which kinetic energy is associated. If one single thing is clear from these definitions, it is that various concepts of electricity must be clarified individually.

Charge

One of the earliest demonstrations of electricity involved rubbing glass or amber until these substances, so activated by friction, had the property of attracting other substances. The glass, or amber, was said to be charged. The early Greeks made this observation. The Greek for amber is "electron," thus the term "electricity." It was also noted that two activated, or charged, amber rods did not attract each other. To the contrary, they repelled each other. Thus one conceives of electrical charges of two types, **positive** and **negative**. Substances with a similar charge repel one another; those with unlike charges attract one another.

It is easy to demonstrate that the force of this attraction or repulsion is a function of distance; the force decreases with the square of the distance. Thus if the force is determined when the particles are 2 units apart and again when they are 6 units apart, the force of attraction or repulsion will be found to have decreased to one ninth $\dfrac{1}{\left(\dfrac{6}{2}\right)^2}$ of the original value.

Another factor which influences the force of attraction or repulsion is the medium separating the charged particles. In air, for example, particles of opposite charge would attract each other with far more force than in water.

Coulomb's Law

The French physicist, Charles A. deCoulomb, formulated an equation which expresses the relationship between charged particles in terms of their distance and the medium. This equation, now known as a law, states that the force exerted between two electric charges is directly proportional to the product of the charges and inversely proportional to the square of the distance between them. Or:

$$F = \frac{q_1 \cdot q_2}{x^2 \cdot K}$$

where: F = force in dynes
q = magnitude of the charge of the particle in statcoulombs
x = the distance between the particles in centimeters
K = the dielectric constant

A **dyne,** it will be recalled, is defined as the force which gives a mass of one gram an acceleration of one centimeter per second. A **statcoulomb** is the unit of charge in the electrostatic system of units. It may be defined as that charge which placed one centimeter from an equal and like charge in a vacuum will repel it with a force of one dyne. This definition, obviously, is derived from Coulomb's law. The **dielectric constant** is a value which expresses the characteristic of the medium in so far as attraction or repulsion between charged particles is concerned. The standard of comparison is a vacuum which has been assigned a value of 1. Air has a value of 1.00059. Distilled water, by comparison, has a dielectric constant of 80.

Potential

One of the most used concepts in cell physiology is that of electrical potential. One speaks of resting potentials, action potentials, and injury potentials, therefore a clear understanding of electrical potential is essential.

As expressed by Coulomb's law, there is a force exerted by two charged particles. If the particles are of opposite charge and attract each other, then, in order to keep them apart, or to separate them, work must be done. Conversely, if they have the same charge, they will repel one another and work must be expended to bring them together. The electrical potential is a measure of work done in moving such charges. By definition, the electrical potential at any distance between two charged particles, is the work done to bring one of the particles from infinity, i.e., a great distance, to a position representing the desired distance from the other. If the two particles have the same charge they will tend to repel one another, therefore, the work done is positive and the potential is positive. On the other hand, if the particles attract one another, the potential is negative. In actual units, the potential is expressed in **statvolts** which is the work done in ergs in bringing a particle of one statcoulomb charge from infinity to a point closer to another charged particle of the same sign. An **erg,** it should be remembered, is the amount of work done when a force of one dyne is exerted on a body and the body is displaced one centimeter in the direction of that force. An erg, then, is a dyne-centimeter.

Actually, in order to determine the potential all one need know is the charge of both particles, and the distance between them. Then:

$$V = \frac{q}{r}$$

where: r = the distance in centimeters between the particles
$\quad\quad\quad q$ = the charge in statcoulombs of the stationary particle
$\quad\quad\quad V$ = the potential in statvolts, that is, ergs per statcoulomb charge of the moved particle

If instead of moving the particle from infinity to a specific point within an electrical field, the particle is moved from a point already in the field to another point in the field then the work done is termed the **potential difference**. From the equation used for calculating potential it follows that:

$$\text{Potential Difference} = \frac{q}{r_1} - \frac{q}{r_2}$$

where: q = the charge in statcoulombs of the field
$\quad\quad\quad r$ = the distance in centimeters between the charged particles

The potential difference, therefore, would be expressed in statcoulombs, that is, ergs per statcoulomb charge of the moved particle.

Field

It has been mentioned that the force of attraction or repulsion of charged particles varies inversely with the square of the distance. It does not matter whether one particle is to the right or left, above or below the other. If the distance between the two is constant, the force will be constant. Accordingly, about each particle there exists a sphere of influence, called the electrical field.

ELECTRICAL CURRENT

The terms volts, ohms, amperes and galvanometer are used so commonly in the field of electricity that a word concerning the origin of these terms is appropriate.

Luigi Galvani was an eighteenth-century Italian physician. He made the observation that when two different metals touched the sciatic nerve of a frog, contraction in the leg muscle occurred every time he brought the other ends of the metals together. The reaction

of the muscle was a response to an electrical current, often called a galvanic current or stimulus.

Another Italian of the same period, Alessandro Volta, a physicist, invented the first electric cell. Again two different metals were used. Volta found that when he placed two metal plates in a dilute acid a current would flow along a wire which connected the plates. He then connected a number of such cells in series to create a battery.

It was only a few years after these demonstrations that G. S. Ohm, a German, formulated a theory for the flow of electricity based upon the principles which govern the flow of heat. He thus concluded that the current is proportional to the potential difference.

A French physicist by the name of André Marie Ampère also carried out original work in this field which he termed electrodynamics. To honor his contributions an international committee in 1881 termed the unit of current the ampere. It was at this same meeting that the volt and ohm were defined.

Magnetic Effects

The early investigators clearly demonstrated that under appropriate conditions a flow of charge could be produced. At that time it was assumed that positive charges flow through a wire from the positive pole to the negative pole. Later it was shown that negative charges, that is, electrons, move in the opposite direction. But whatever the direction, some means was clearly needed to detect the existence of an electrical current. It was the discovery that the passage of a current produced a magnetic effect that made this possible. Thus, if a wire conducting a current is held over and parallel to a compass needle, the needle moves and assumes a position at right angles to the wire. It is easy to demonstrate that there is a **magnetic field** surrounding every current of electricity.

Galvanometer

If no current flows in a wire, there is no magnetic effect. On the other hand the stronger the current, the greater the magnetic field; this is the principle of the galvanometer. In such an instrument, wire is wound many times in the shape of a circle or flat rectangle; the more wires the stronger the magnetic effect. A compass needle is

then mounted so that it can pivot at the center of the coil of wire. When a current passes through the coil, the needle will be deflected proportionately to the strength of the current.

Ampere

The term **current** has been used without defining it. A current is the quantity of electric charge transferred past a point per unit time. The unit of current is the ampere. It was originally derived in terms of the magnetic effect. Actually, an ampere is defined as being one tenth of an abampere. An abampere is the current which, flowing in a single turn of circular wire of one centimeter radius, exerts a force of 2 pi dynes on a unit magnetic pole placed at the center. This is a rather involved definition and for practical purposes the ampere may be defined as a flow of a charge of 1 coulomb per second.

Coulomb

The statcoulomb has been defined. Before the coulomb is discussed, it must be understood that there are at least three sets of units used in electricity. There is the **electrostatic system** for measuring charge and related phenomena. In this system statcoulombs and statvolts were discussed. Electrical phenomena may also be defined in terms of the **magnetic effects** of the current, as was done in the case of abampere. And still a third method is to use the so-called **practical units.** One can always tell which system is being used by noting the prefix. Stat- indicates electrostatic units, ab- electromagnetic, and if no prefix is used then practical units are being employed. In this system the coulomb is defined as the quantity of electric current that passes a given point in a conductor in a given time. In other words:

$$\text{coulombs} = \text{amperes} \times \text{seconds}$$

One coulomb, then, is conducted by a current of one ampere in one second.

Volt

The volt is a measure of potential difference. In the electrostatic system this unit is termed the statvolt. In practical units the volt is

defined in terms of resistance and current. Thus before the volt is discussed further, the concept of resistance must be introduced.

Ohm

As has already been mentioned, Ohm showed that the current flowing in a conductor is proportional to the potential difference. If the potential difference at the two ends of a metal conductor remains constant, the current will be constant. When the potential difference is increased, the current increases. On the other hand, Ohm also showed for any potential difference the current varied with the conductor used. A poor conductor may be thought of as resisting the flow of current. Thus various conductors may be classified according to their resistance to the flow. When the three factors, potential difference, current, and resistance are put together, it is found that:

$$I = \frac{V}{R}$$

where: I = current expressed in amperes
V = potential difference expressed in volts
R = resistance of the conductor expressed in volts

By definition, then, a conductor with a resistance of one ohm conducts a current of one ampere when the potential difference is one volt. This may be expressed as:

$$R = \frac{V}{I}$$
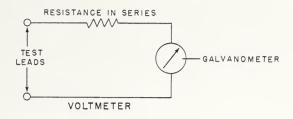
A volt may now be defined, in practical units, as the electromotive force (potential difference) which, when steadily applied to a conductor whose resistance is one ohm, will produce a current of one ampere. This may be expressed as:
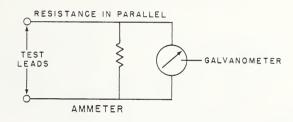
$$V = RI$$

Voltmeter

A galvanometer, as discussed above, measures current. Such an instrument may also be used to measure voltage simply by adding a high-resistance coil and placing the galvanometer in parallel with

the current (Fig. 18.1). Thus, very simple modern instruments can be used to measure voltage, current, and resistance. By the use of a series of shunts and resistances the instrument is made multirange.
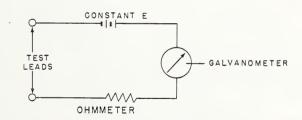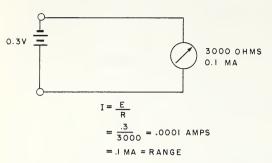
Fig. 18.1. Basic Circuits for Using a Galvanometer for Measuring Voltage, Current, and Resistance.

When the instrument is used to measure current, it is called an **ammeter,** to measure resistance an **ohmmeter.**

A few examples and calculations will serve to clarify these relationships. Figure 18.2 shows a simple galvanometer which has an internal resistance of 3000 ohms. It is found that with a voltage of 0.3 volts and therefore a current of 0.0001 amperes, or 0.1 milli-

amperes (MA), the needle is deflected full scale. If this instrument is connected in series with a circuit, it will serve as an ammeter with a range of 0–0.1MA. As shown in Fig. 18.2 a shunt with a resistance lower than 3,000 ohms will increase the range. For example, in order to make the range 0–1MA it is necessary to shunt 0.9MA $(1 - 0.1)$ before it reaches the meter.

$$I = \frac{E}{R}$$

$$= \frac{.3}{3000} = .0001 \text{ AMPS}$$

$$= .1 \text{ MA} = \text{RANGE}$$

R= 333 OHMS

3000 OHMS
I MA

$$I - 0.1 = 0.9 \text{ MA}$$

$$.0009 = \frac{.3}{R}$$

$$R \equiv 333 \text{ OHMS}$$

$$\text{RANGE} = I \text{ MA}$$
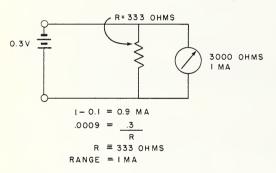
Fig. 18.2. Basic Circuit for an Ammeter. The addition of 333-ohm resistor extends the range from 0.1 MA to 1 MA.

Since $R = \dfrac{E}{I}$

$$= \frac{0.3}{.0009}$$

$$R = 333 \text{ ohms}$$

In a similar manner it is calculated that a shunt with a resistance of only 33.3 ohms will provide an ammeter with a range of 0–10MA.

The galvanometer may be used as a voltmeter if it is placed in parallel with the circuit. This particular voltmeter would then have a full scale range of 0–0.3V. To extend the range to 5 volts an appropriate resistance must be added in series (Fig. 18.3).

Again, $R = \dfrac{E}{I}$

$$R + 3,000 = \frac{5}{.0001}$$

$$R = 47,000 \text{ ohms}$$

In other words, a resistance of 47,000 ohms must be placed in series with the original internal resistance of 3,000 ohms to provide a volt-

meter with a range of 0–5V. It should be added that the ideal volt-meter is one that draws no current, while the ideal ammeter is one requiring no voltage.

And finally, the galvanometer may be converted to an ohmmeter simply by providing a small dry battery as a source of d.c. voltage (Fig. 18.1). The scale is calibrated to read the un-known resistance directly. In this arrangement only resistances below that of the internal resistance may be determined. The circuit for measuring re-sistances over a wide range is shown in Fig. 18.4. In this case a high-resistance voltmeter is used to measure the volt-age drop across a refer-ence resistor. The range of such an ohmmeter is

0.3 V

3000 OHMS
0.1 MA

R = 47,000 OHMS

5 V

3000 OHMS
0.1 MA

$$R + 3000 = \frac{5}{100 \times 10^{-6}}$$

$$R = 47,000 \text{ OHMS}$$
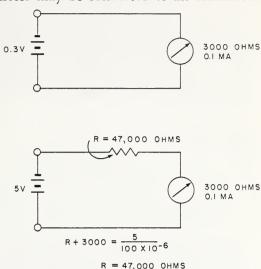
Fig. 18.3.   Basic Circuit for a Voltmeter. The addi-tion of 47,000-ohm resistor extends the range from 0.3 volts to 5 volts.

limited only by the resistance of the meter itself. This resistance is usually high enough so that with the use of a variable resistor the range of the ohmmeter extends from ohms to megohms. In this in-strument the current generated by the battery must be great enough to overcome the resistance of the total circuit, which includes the un-known resistor and the reference resistor. There-fore, a battery of suitable voltage should be se-lected. Generally, a 3-volt battery suffices for an ohmmeter with a maxi-mum range of 100,000 ohms. A 6-volt battery is used for higher ranges.

ZERO ADJ.

R_1

HIGH
RESISTANCE

R_1 = REFERENCE RESISTOR
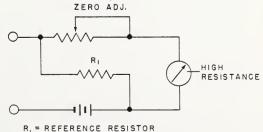
Fig. 18.4.   Basic Circuit for an Ohmmeter.   By changing the reference resistor the range can be greatly extended.

ELECTRICAL STIMULATION

In the physiological laboratory, cells, tissues, and organs are stimulated and their response to the stimulation studied. A stimulus may be chemical, osmotic, mechanical or electrical. For example a nerve will respond when a drop of acid is placed on it, or when it is dipped in a concentrated salt solution, or when it is pinched. It will also respond to an electrical shock. Since the shock can be controlled as to amplitude, frequency and duration far more easily than the other types of stimuli, it is usually the one chosen for laboratory experimentation.

Direct Current

A direct current is one in which the charged particles always flow in the same direction. Thus if the terminals of a battery are connected, there will be a flow in one direction and it will be continuous for the life of the battery. This was the type of current produced by Galvani, and it is often referred to as a **galvanic** current.

Electronic stimulators which are now widely used in physiological research do not emit a continuous current, but rather pulses of direct current. Since flow is always in the same direction in each pulse, the output from such a stimulator is said to be **monophasic.** Figure 18.5 portrays various important attributes of this type of stimulus. In the first place the wave is seen to be square. This means that the pulse reaches its particular magnitude instantaneously, holds it for a specific duration and then returns to zero voltage instantaneously. Only the better stimulators approach this ideal. The three param-
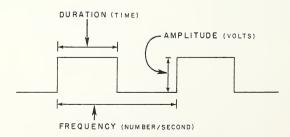
Fig. 18.5.   The Duration, Amplitude, and Frequency of Square Waves.

eters of the pulse are the amplitude measured in volts, the duration measured in milliseconds, and the frequency expressed in pulses per second.

Alternating Current

Instead of having the charged particles flow only in one direction, it is possible to make them flow first in one direction and then in the opposite direction; such a current is said to alternate. Again the shape of the wave may be square, but more commonly it is sinusoidal (Fig. 18.6). Each complete biphasic wave is said to constitute one cycle. In the usual house current there are 60 cycles per second.

Alternating current is rarely used to stimulate tissues but it does have important usage in

Fig. 18.6.   Alternating Current.

the clinic, namely in diathermy treatment. Diathermy means "a heating through." The tissues under treatment resist the passage of the current and therefore become heated. There is thus penetration of the heat, in effect, a heating through. When an alternating current of high frequency is used there is no stimulation to the tissue in the sense of depolarization, but there is a heating effect. Thus, this is the type of current used in diathermy treatment.

Induced Current

In many student laboratories, electrical stimuli are produced by means of an **inductorium**. This type of stimulator consists of two coils of wire, one of which can be moved in relation to the other. A direct current flows through the so-called primary coil and, as a result, there is an induced current in the secondary coil. It will be recalled that there is always a magnetic field surrounding a wire conducting a current. This field represents a source of potential energy. Energy is stored in the field. In short, there has been a change of electrical energy to magnetic energy. This represents a back electromotive force which, in time, comes to balance exactly the induced electromotive force. The only way to alter the induced energy is to vary the current in the wire. Thus, in the inductorium, the voltage

in the secondary coil is proportional to the **rate of change** of the current in the primary coil.

The unit of inductance is the **henry** and its definition will make these relationships clear. A coil is said to have an inductance of 1 henry when a current flowing through it and changing at a rate of 1 ampere per second produces a back voltage of 1 volt.

A typical inductorium is shown in Fig. 18.7. The wave form is biphasic, the direction of the waves being opposite on the "make" and the "break." The make wave results from the introduction of
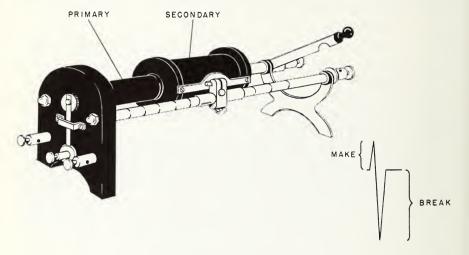
Fig. 18.7.   An Inductorium. The amplitude of induced stimulus is greater on the "break" than on the "make."

the current into the primary coil. The break wave occurs when the current is stopped. It will be noted that the magnitude of the output from the secondary coil is much greater on the break than on the make. This follows from the fact that the rate of change of current in the primary coil is much faster when it is stopped than when it is started. To put it another way, the current builds up more slowly than it cuts off.

That there are two stimuli of different magnitudes from an inductorium immediately becomes apparent in a simple experiment in which a muscle is stimulated. If the two coils are separated sufficiently so as to produce a minimal stimulus, the muscle will be seen

to respond once when the current is introduced, and to respond a second time and with greater force when the current is stopped.

SUMMARY

Electricity consists of charged particles termed **electrons** which are negative and **protons** which are positive. Particles of similar charge repel one another; those of opposite charge attract. According to **Coulomb's law** the force exerted between two electric charges is directly proportional to the product of the charges and inversely proportional to the square of the distance between them. The force is measured in **dynes** and the charge expressed in **statcoulombs.** A statcoulomb is defined as that charge which, when placed one centimeter from an equal and like charge in a vacuum, will repel it with a force of one dyne. The **dielectric constant** is a value that expresses the characteristic of the medium in which charged particles attract or repel one another.

The work required to change the distance between two charged particles is termed the **potential difference.** The work is expressed in **ergs;** the potential, in **statvolts** which is the work done in bringing a particle of one statcoulomb charge from infinity to a point closer to another charged particle of the same sign.

When charged particles flow in a conductor, a **magnetic field** is created. Thus, the field is proportional to the current, and a compass needle will be deflected in accordance with the current. This is the basis of the **galvanometer.** The **ampere** is the unit of current and is defined, in practical units, as a flow of charge of 1 coulomb per second. A **volt** is the electromotive force which, steadily applied to a conductor whose resistance is one ohm, will produce a current of one ampere. An **ohm** is the resistance of a conductor that will carry a current of one ampere when the potential difference is one volt. The galvanometer may be used with appropriate resistances as an ammeter to measure current, and as a voltmeter to measure voltage. By the addition of a small dry battery as a source of d.c. voltage it may be used as an ohmmeter to determine resistance.

Electrical stimulation is most often used to activate living cells. Electronic stimulators produce pulses of **direct current. Alternating**

current of high frequency produces penetrating heat in tissues. The wave form from the output of the secondary coil of an **inductorium** is biphasic. The "make" wave is generally of lesser amplitude than is the "break" wave. The unit of inductance is the **henry.** A coil is said to have an inductance of 1 henry when a current flowing through it and changing at the rate of 1 ampere per second produces a back voltage of one volt.

Problems

1. What is the force of attraction between two particles in water 8 cm apart, one charged with +6 and the other with −6 coulombs?
2. If the electromotive force is 120 volts and the resistance 2,000 ohms, what will be the current?
3. A voltmeter with an internal resistance of 50,000 ohms has a range of 10 volts. How could the range be increased to 50 volts?
4. An ammeter has an internal resistance of 5,000 ohms and a range of 1 MA. How could the range be increased to 10 MA?
5. Explain, with the aid of a drawing, how a multirange ohmmeter might be constructed.
6. Why do the "make" and "break" shocks from an inductorium vary in amplitude?

Additional Reading

Catton, W. T., "Physical Methods in Physiology," Philosophical Library, New York, N. Y., 1957.

Mendenhall, C. E., Eve, A. S., Keys, D. A., and Sutton, R. M., "College Physics," D. C. Heath and Co., Boston, Mass., 1944.

Randall, J. E., "Elements of Biophysics," The Year Book Publishers, Inc., Chicago, Ill., 1958.

Stacy, R. W., "Biological and Medical Electronics," McGraw-Hill Book Co., New York, N. Y., 1960.

"The Radio Amateur's Handbook," The Rumford Press, Concord, New Hampshire, 1958.

CHAPTER 19

METHODOLOGY

A SCIENCE is an accumulation of knowledge and the acquisition of that knowledge depends upon observation. Unaided powers of observation are limited. Methods are constantly being developed to enhance the ability to observe. The growth of a science inevitably is limited by, and keeps step with, the rate of development of its methodology. Not only is the growth of a science dependent in this way, the comprehension of that knowledge requires an understanding of the methods used in acquiring the basic data. In this chapter an attempt is made to acquaint the reader with methodology essential to physiology.

MICROSCOPY

One of the prime purposes of any instrument is to increase the efficiency of our own senses. The microscope, of course, does just that for the sense of vision. It has been said, with obvious validity, that without microscopes there would be no cytology. Knowledge of the internal structure and organization of the cell became possible only after the invention of the microscope. The first, truly functional microscope was made by a Dutchman, Zacharias Janssen in about 1590. Until the beginning of the 20th century there was little improvement in microscopes, but since that time progress has been very rapid and significant.

Limitations of Microscopy

The purpose of using a microscope is to see objects that are too small to be seen with the unaided eye. What limits this size? The

ultimate limitation is the wavelength of the light used to form the image of the object. This limitation is termed the **resolution** of the microscope. It can be shown that the smallest detail that can be imaged by a microscope is about a half the wavelength of the light with which it is seen. The shortest wavelength of visible light is about 4000 angstrom units (Å). One angstrom unit equals 1×10^{-7} millimeter. Light microscopes of this resolution have been developed. The only way to visualize still smaller objects clearly is to use illumination with a wavelength shorter than that of visible light.

Phase Contrast Microscopy

Size is but one of the limiting factors in determining what is seen with a microscope. Another important consideration is the contrast the object makes with its surroundings. The phase contrast microscope improves contrast. When rays of light pass from one medium into another at an angle other than 90 degrees, they are bent, that is refracted, if the speed of transmission of light in the two media differs. A cell, for example, is not homogeneous, therefore, different areas refract light to a greater or lesser degree. Deflected rays of light are out of phase with the light directly transmitted and as a result there is produced an image of strong contrasts.

The phase contrast microscope permits observation of unstained, living cells, which, of course, has great advantage since staining sometimes alters the cell structure. But it is also useful with lightly stained cells since the stained objects will be presented with far greater contrast than under normal illumination.

Ultraviolet Light Microscopy

As has been mentioned, shorter wavelenghts permit higher resolution. Ultraviolet light has a wavelength shorter than that of visible light. However, normal optical glass is unsuitable because it is opaque to ultraviolet light. Lenses have been developed which permit the passage of wavelengths as short as about 2400 Å. Quartz is usually used for this purpose; CaF_2 is also used. Lithium fluoride is transparent to wavelengths even below 2000 Å and thus is very valuable for this type of microscopy.

The ultraviolet microscope has proved especially useful in the study of chromosomes and other minute constituents of the cell. Liv-

ing cells may be studied with this technique but, unfortunately, the ultraviolet light has a destructive influence on living tissue.

Electron Microscopy

The electron microscope has a resolving power more than one hundred times greater than that attainable with visible light. In this technique, beams of electrons are used in place of light rays. Electrons are charged particles that can be focused by causing them to pass through appropriate electrostatic and magnetic fields. With the electron microscope objects with a diameter as small as 10 Å, or even somewhat less, may be resolved. Since the average atom has a diameter of about 2 Å, it can be appreciated that the electron microscope is an invaluable tool.

The limitations of the electron microscope are in the sample preparation. In the first place, the sample must be viewed in a vacuum which restricts the use of living tissue. And secondly, to take advantage of the great resolving power, the sample must be ultra-thin. This is a difficult task. However, new techniques of tissue preparation are being developed and the use of the electron microscope is of rapidly increasing importance to the science of cytology.

ANALYTICAL METHODS

Colorimetry

A simple analytical method involves the comparison of the light transmitted by an unknown solution with that transmitted by a standard solution. A colorimeter, then, is an instrument used to make this comparison.

Colorimetry is based upon two laws: 1) Bouguer's and 2) Beer's. **Bouguer's law** states that when a ray of monochromatic light enters an absorbing medium, its intensity decreases exponentially with an increase in the thickness of the medium traversed. **Beer's law** states that the intensity of a ray of monochromatic light decreases exponentially as the concentration of the absorbing material increases. When these two principles are combined into what is known as the Bouguer-Beer law, it is seen that the following relationship holds:

$$\frac{c_1}{c_2} = \frac{d_2}{d_1}$$

where c equals the concentration of the solution, and d the depth through which the light is transmitted. Clearly, if one knows the concentration of the standard solution and the depth the light is transmitted in both solutions when they compare, then the concentration of the unknown solution may be easily calculated (Fig. 19.1).
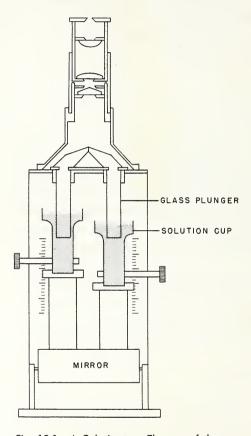
Fig. 19.1. A Colorimeter. The cup of the unknown solution is moved up or down until the light transmitted by the two solutions appears equally bright. The depth of the solution between the bottom of the cup and the glass plunger is read directly for the two solutions. Since the concentration of the standard is known, the concentration of the unknown may be readily calculated (see text).

Photometry

Instead of comparing an unknown solution with a standard, it is possible to determine directly the amount of light transmitted by the solution. An instrument that accomplishes this is termed a **photometer.** A photoelectric cell is usually used to determine the intensity of the transmitted light. Such a cell emits an electrical current that is proportional to the light intensity. Consequently, the resulting current measures the light intensity. Obviously, a properly constructed electrophotometer gives far greater precision than is possible using a comparison colorimeter.

Spectrophotometry

In photometry the amount of light transmitted by a solution is measured. Spectrophotometry differs in that a specific wavelength of light transmitted by the solution is measured. In other words, a spectrophotometer is a filter photoelectric photometer with a range including many nearly monochromatic bands of light. But instead of changing filters to obtain a specific band of light, the spectrophotometer employs a prism so that by turning a knob a particular band of light is obtained. Not only is it easier to change wavelengths this way, but the purity of the light energy is much greater. Accordingly, a spectrophotometer makes it possible to determine the concentrations of components of mixtures easily and precisely.

Flame Photometry

The flame photometer provided a tremendous impetus to ion research, especially work involving potassium, sodium, and calcium because it permits rapid, simple, and accurate analysis. Other analytical procedures for these ions are extremely time consuming.

The principle of the flame photometer is that the intensity of energy emitted by a flame is proportional to the concentration of the ion that excites it. In practice the ion is atomized into the flame and there results a characteristic wavelength of energy which is then measured. Over 30 elements have been found to excite a flame and therefore can be analyzed with this method. The greatest use in the physiological laboratory, however, has been in the analysis of sodium, potassium, and calcium.

Radioisotope Techniques

There are many methods for measuring radioactivity. In the biological laboratory one most often determines the beta or gamma radiation by the use of either a Geiger-Muller tube (G-M), or a scintillation detector.

The **G-M tube** contains two electrodes and is filled with a gas, usually argon or neon. A high-voltage power supply furnishes about 1000 volts between the electrodes. When radiation enters the tube, the gas is ionized; because of the high-voltage, ionization is rapid and an electrical pulse results. The discharge of the tube caused by ionization terminates when enough positive ions collect around the negative electrode to reduce the electrical field strength so that no further ionization occurs. Most G-M tubes have a quenching gas to hasten this process and thus prepare the tube to respond to further radiation. The entire sequence must occur in a fraction of a millisecond.

A **scintillation detector** is more sensitive than a G-M tube and is most generally used in the detection of gamma rays, though with special crystals it can be used for betas. The principle of the scintillation counter is the emission of light by certain crystals as a result of absorbing radiation. The intensity of the scintillation is proportional to the energy of radiation. The number of pulses is proportional to the frequency of the radiation. Consequently, with a scintillation detector it is possible to determine the energy of radiation as well as the frequency.

With either type of detector the frequency of the radiation is determined, or counted, by either a scaler or a count rate meter. A **scaler** is an electronic counting apparatus. It counts the number of pulses coming from the G-M tube or the scintillation detector. Thus, if one leaves the instrument on for precisely one minute, the number recorded in that time will indicate the pulses, or counts per minute. A **count rate meter** has an integrating mechanism so as to convert counts into counts per minute which are then read directly. The advantage of a count rate meter is that it can be used to activate a recorder so as to obtain an automatic record of the isotope intensity, or a record of rapid changes in activity.

A **spectrometer** is often used between the scintillation detector and the scaler, or rate meter, to discriminate between pulses of dif-

ferent amplitudes (Fig. 19.2). In this way background activity may be reduced, or individual isotopes in a mixture determined. A spectrometer functions by permitting pulses of a predetermined amplitude to pass and therefore to be counted. Accordingly if there is a mixture of two isotopes and their gamma rays have different energy values, either one or both may be counted by using a spectrometer.

Because isotopes of the same element are chemically indistinguishable, radioactive isotopes find wide application in physiological research. Such radioactive substances, by the above described means, may be detected with great accuracy and in extremely small quantities.

Autoradiography

A method that combines histological and radioisotope techniques is called autoradiography. In this procedure a radioactive isotope is introduced into a cell and then the cell is brought into contact with a sensitive emulsion which is exposed by the radioactivity. The technique and the emulsions have been so greatly improved in recent years that truly excellent resolution is now possible. By the use of autoradiography radioactive atoms may be localized with great precision in various cell structures. Because the method is so sensitive, extremely low doses of the isotope may be used. This method has great promise of contributing significantly to the knowledge of cellular metabolism.

Mass Spectrometry

Because of the complexity of the method and the expense of the instrument, mass spectrometry has not, as yet, become widely used in physiological work. The mass spectrometer sorts out ions according to the ratio of their charge to their mass. These ions are then collected and their intensity measured. In this way the amount of any ion in a complex mixture may be determined with great accuracy even when the ion concentration is very low.

Chromatography

It has been found that when a mixture of substances in solution pass over or through an adsorbent, the substances separate into spe-
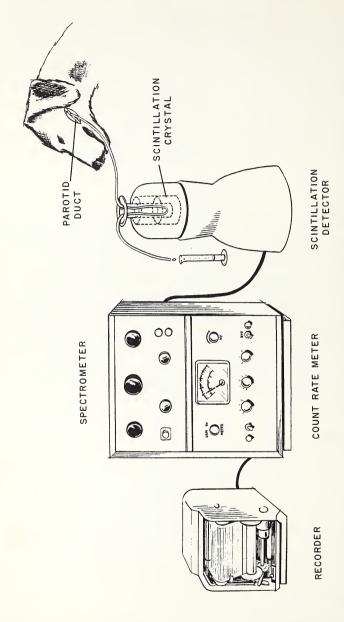
Fig. 19.2. Instrumentation Used in the Author's Laboratory to Study the Physiology of Salivation. The activity of the isotope in the saliva is detected by the scintillation crystal. The resulting impulses are then led into a spectrometer, to a count rate meter, and to the recorder.

cific zones dependent upon the rate at which the substance moves through the adsorbent. This is an excellent method for separating organic compounds. Chromatographic paper has been developed which is used as the adsorbent. After the separation has occurred the substances are usually identified and quantitatively determined. An excellent way of doing this combines chromatographic and radio-isotope techniques. For example, an organic compound is prepared with a radioisotope in its molecule. It is separated from the mixture and identified by chromatography. Radioisotope techniques can then be used to locate the radioactive element.

Warburg Technique

One of the most widely used procedures in cell physiology is the Warburg technique. The apparatus used was originally developed by Barcroft and Haldane, but it was adopted by Warburg, in 1912, for metabolic studies. Essentially this is a procedure for determining the rate of oxygen utilization of cells, or small segments of tissue. The cells to be studied are placed in an appropriate medium in a small glass vessel. The vessel is generally filled with pure oxygen and a substance used to remove the carbon dioxide resulting from meta-bolic processes. The vessel is submerged in a constant water bath and agitated. The amount of oxygen utilized is determined mano-metrically. It is also possible, in specially designed Warburg ap-paratus to determine the carbon dioxide production as well as the oxygen utilization.

The Warburg technique has been developed to the point where it is also possible to follow the rate of utilization of a foodstuff or sub-strate which has been added to the medium in which the tissue lies. This technique permits the *in vitro* study of metabolic transforma-tions and is directly responsible for the considerable knowledge that now exists concerning enzymatic systems.

Electrophoresis

Proteins migrate in an electric field. This is true, of course, only at a pH other than that of the isoelectric point. At the isoelectric point there is no net charge and therefore no migration in an electric field. The rate of migration is a function of the net charge. The net charge can be changed by altering the pH of the solution. Accord-

ingly, the rate of migration can be determined at different pH values. The rate of migration, termed the electrophoretic mobility is generally expressed in square centimeters per volt per second. The electrophoretic technique, then, involves adding acid or base to the solution containing protein to alter the pH. The rate of migration is measured and the results plotted (Fig. 19.3). The isoelectric point
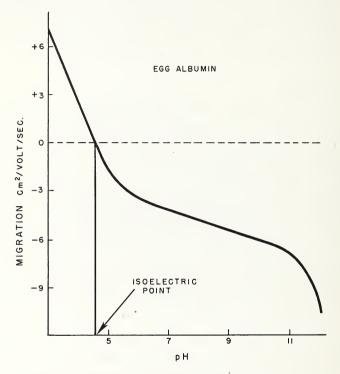
Fig. 19.3.   The Rate of Migration of Egg Albumin at Different pH Values. Zero migration represents the isoelectric point.

becomes readily apparent. The so-called titration curve obtained in this way differs with each protein. This method, then, is useful in determining the purity of individual proteins and it is also used to identify specific proteins in mixtures.

Instead of determining the rate of migration, it is usually the practice to determine the refractive index of various portions of the protein solution. This provides an indication of the distribution and concentration of specific proteins. Instruments have been developed

which determine the refractive index and record the results in the form of peaks. A single peak indicates a solution of but one protein. Multiple peaks and uneven curves indicate several proteins in a mixture.

Ultracentrifugation

The ultracentrifuge, as the term indicates, is a centrifuge capable of attaining very high speeds, and therefore, of producing very great centrifugal forces. Speeds as high as 80,000 rpm have been attained. At these high speeds forces approaching 500,000 times gravity result. Such force is quite capable of causing large molecules in solution to be sedimented. Accordingly, ultracentrifugation has been used, with notable success, to determine the molecular weight of proteins.

The rate of movement of the protein molecules in the ultracentrifuge depends upon the centrifugal force, the size of the molecule, and the viscosity of the solvent. Taking these factors into consideration, the following equation is used to determine molecular weight:

$$M = \frac{RTs}{D(1 - Vp)}$$

where: M = molecular weight
T = absolute temperature
R = gas constant
s = sedimentation constant
D = diffusion constant
V = partial specific volume of the protein
p = density of the solvent

Histochemical Techniques

Considerable knowledge of the morphology of the cell has been obtained by use of the usual histological procedures of fixing the tissue, slicing ultra-thin, and then staining. But newer techniques are now available which permit the cellular biochemist to study intracellular metabolism. In essence the method consists of treating tissue sections with chemicals that react with certain intracellular substances to produce a specific color. Not only is it possible to ascertain by this method whether or not a particular compound is present in the cell, but the technique has been refined to the point that one can determine the relative concentration and distribution of the substance.

RECORDING METHODS

Instruments not only aid our senses and provide means of analysis, but they also are capable of providing a permanent record of data. It has already been mentioned that in radioisotope work a recorder, generally an ink-writing recorder can be activated by a count rate meter to make a permanent record of the counts per minute. Not only does this provide an automatic and permanent record, but there are cases in which the values change too rapidly to be observed in any other way. Other types of recording devices used in physiology will be briefly described.

Kymograph

In 1847 Ludwig, a German physiologist, invented the kymograph for recording physiological phenomena; it is still widely used today. As shown in Fig. 19.4, the kymograph consists of a rotating drum on which is placed smoked paper. The physiological event under study is caused to activate pointers which make a tracing on the smoked paper. When the record is complete, it is fixed by treating it with shellac. The rate of rotation of the drum can be varied, thus if there is rapid change, a fast rotation is used, if the change is slow, a slower rotation is selected. For most records a timing device is also used.

The kymograhic method of recording is an exceedingly simple procedure which does suffice for many types of student experiments. In addition much outstanding research has resulted from its use. However, there are serious limitations, and far superior methods are now available. The kymograph can be used to record muscle contraction, blood pressure, and respiration. As indicated in Fig. 19.4, the muscle is attached to a lever which has a point that makes a tracing on the smoked paper. For blood pressure, a lever is attached to a float that rides on a column of mercury. As the pressure changes, the mercury rises or falls and so does the float thus making a proportional tracing on the smoked paper. For respiration an air-chamber is used which is covered with a rubber diaphragm. During expiration the diaphragm bulges and the lever attached to it moves thus making a tracing.

The kymographic method has been improved by the development of ink-writing pointers that trace a permanent record on unsmoked

paper. Another improvement involves the use of a special paper upon which a tracing is made by a hot stylus. These procedures eliminate the disadvantages of the smoked paper, but still leave the limitations of a mechanical system.
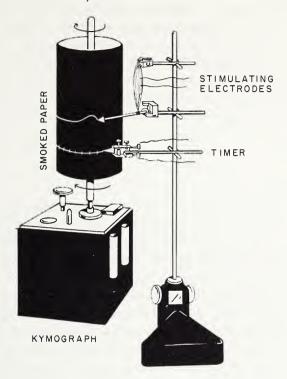
Fig. 19.4. The Kymograph. In this example the kymograph is being used to record muscle activity.

Optical Recording

One of the main drawbacks of the kymographic method is the inertia in the system. This can be avoided by the use of light beams, appropriate optics, and a moving paper coated with a light-sensitive emulsion. For example, to record pressures one can arrange a chamber, the end of which is closed by a rubber diaphragm. A small mirror is attached to the rubber. As the pressure changes, the diaphragm moves and so does the mirror. A beam of light focused on the mirror will then also move. The beam is reflected onto the moving light-sensitive paper and a record is obtained. Such a system gives a much

more faithful record than does one involving the movement of a mechanical lever. Optical recording has wide application. For example, physiological electrical phenomena may be recorded by the use of string galvanometer combined with an optical system. The string is usually made of finely spun silver-coated quartz glass. It is suspended between the poles of an electromagnet. The magnet under the influence of a current creates an electrical field. The potential difference of the physiological event being studied causes a current to flow in the string. The string, therefore, moves in the electrical field in relation to its current. A light source causes the string to cast a shadow on light-sensitive moving paper and thus a record is produced. Such a system has very low inertia and is capable of very rapid responses.

The Oscilloscope

Even less inertia exists in the cathode-ray oscilloscope. Figure 19.5 shows that this instrument consists essentially of a vacuum tube in which a beam of electrons is focused on a fluorescent screen. This beam of charged particles can be deflected by charged plates placed on either side of the beam, in the vertical and in the horizontal planes. The vertical plates cause the beam to move from one side to the other. This is the so-called sweep. The rate of sweep can be varied and if it is fast enough one has the impression of a steady line of light on the screen. Actually, it is sweeping from one side to the other, is then cut off until it returns to the original side and then sweeps again. A potential difference on the horizontal plates will
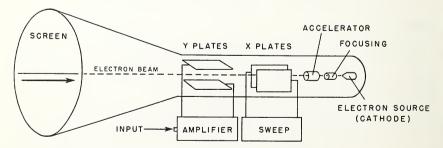
Fig. 19.5. The Oscilloscope. A potential difference between the X-plates causes the electron beam to sweep across the screen. A potential difference between the Y-plates causes the electron beam to move up or down.

cause the beam to move up or down. Thus, if the beam is sweeping from one side to the other and, at the same time, is caused to go up and then down, a curve will be obtained on the screen for a fraction of a second. To make a permanent record one photographs the screen. The oscilloscope thus can be used to record any event that can be converted to a proportional potential difference. For example, electrodes are placed one on the surface of a cell and one within the cell, and these electrodes are attached to the horizontal plates of the oscilloscope. The potential difference between the inside and outside of the cell membrane will deflect the beam of electrons. The screen may be photographed so as to obtain a permanent record.

Ink-Writing Polygraph

If the physiological event under study creates an electrical impulse, or if the event can be converted to an electrical impulse, a recording of the event can be made in a variety of ways. One method is to use the oscilloscope just described. Another procedure involves amplifying the impulse adequately so that a pen may be activated. The movements of the pen are then recorded on moving paper (Fig. 19.6).

The ink-writing polygraph is an excellent instrument both for student experiments as well as for research. Most such instruments have at least three channels each consisting of an amplifier and a pen. There is often also a fourth pen that is used as a timer-indicator. In other words, this fourth pen is activated by a timing device so as to make marks on the paper say every second, or every five seconds. It also makes a mark when the electronic stimulator is used. Thus the exact time of stimulation is automatically indicated and the response automatically recorded.

An ink-writing polygraph has more inertia than an optical system, or the oscilloscope, but it has definite advantages in that it is simple to use and the permanent record is immediately visible. The sole limitation to this type of electronic recorder is the transducer. A **transducer** is an instrument that converts one form of energy into another form. In this case, it converts the energy of the physiological event into an electrical impulse. If a proportional electrical impulse can be obtained, the event can be recorded.

Membrane potentials, of course, need no transducer, merely suffi-

cient amplification. Such physiological events as muscle contraction, blood pressure, respiration, heart sounds, blood flow, oxygen saturation of the blood, and secretion may all be converted to proportional electrical impulses and thus recorded. In addition, by combining electronic recording with many of the methods discussed in this chapter records of cellular activity may also be made.
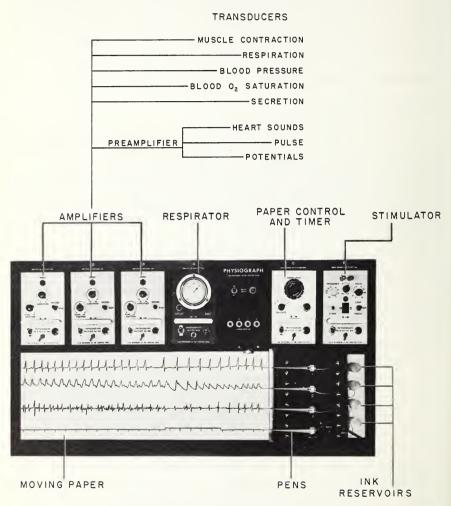
Fig. 19.6. The Polygraph. There are many types of ink-writing polygraphs. The one shown here was developed not only for research but also to serve as standard equipment in the student laboratory. As indicated, by the use of different transducers, a wide variety of physiological events can be recorded.

SUMMARY

The microscope permits objects to be studied that are too small for the unaided eye. The ultimate limitation is not in the optics of the instrument but rather in the wavelength of the light used. The shortest wavelength of visible light is about 4000 Å. Objects with a diameter of about half this length may be visualized. The **phase contrast microscope** gathers rays of light refracted by dense objects out of phase with rays that are transmitted more directly. Accordingly, there is interference which yields strong contrasts. This technique permits unstained, living cells to be studied. To increase the power of resolution of a microscope, ultraviolet light is used, or a beam of electrons. With the **electron microscope** objects with a diameter of only 10 Å may be visualized.

Colorimetry is an analytical method which depends upon the comparison of transmitted light through a standard with that transmitted by an unknown solution. In a **photometer** the intensity of the transmitted light is determined. A **spectrophotometer** measures the intensity of a specific wavelength of light transmitted. **Flame photometry** is based on the principle that the intensity of energy emitted by a flame is proportional to the concentration of the ion that excites it.

Radioactivity may be detected by either a **Geiger-Muller tube,** or a **scintillation detector.** Both convert the energy of radiation into electrical impulses which are then counted by either a **scaler** or a **count rate meter.** Since the pulse amplitude emanating from the scintillation detector is proportional to the energy of the radiation, a **spectrometer** may be used to determine that energy value or to discriminate between pulses of different amplitudes. A **recorder** may be activated by a count rate meter to obtain an automatic, permanent record.

The **mass spectrometer** sorts out ions according to the ratio of their charge to their mass. The ions are then collected and their intensity measured. In **chromatography,** compounds are separated by passing the solution through an adsorbent. They may then be determined and their concentration measured. The Warburg technique is used to determine the rate of oxygen utilization by cells and tissues. **Electrophoresis** permits specific proteins to be identified.

The **kymograph** is a recording method which uses a rotating drum

upon which moving levers make tracings. In **optical recording** there is less inertia because a beam of light is caused to move in response to the event being studied and the light is then reflected onto moving light-sensitive paper. The **oscilloscope** employs a beam of electrons which activates a fluorescent screen. The beam is caused to sweep across the screen and is deflected in the vertical plane in response to the physiological event. By suitable amplification an electrical impulse can be made to activate a pen which writes on moving paper. **Transducers** are used to convert the energy of the physiological event into an electrical impulse. Two or more pens are combined for simultaneous recording in the **polygraph.**

Problems

1. Outline the advantages of the phase contrast microscope.
2. Why may smaller objects be visualized with the ultraviolet light and electron microscopes than with a visual light microscope?
3. Compound XYZ is highly soluble and gives a characteristic color to the solution. Explain at least two methods that could be used to determine the concentration of this substance in solution.
4. Isotope A has a half-life of 15 hours, a beta radiation of .78 Mev, and a gamma radiation of 1.6 Mev. Isotope B has a half-life of 8 days, a beta radiation of .61 Mev, and a gamma radiation of .86 Mev. Isotope C has a half-life of 7 days, a beta radiation of .72 Mev, and no gamma. Explain how each may be counted when in solution alone. Explain how each may be counted in the following combinations: A + B, A + C, C + B, A + B + C.
5. What are the advantages of the ink-writing polygraph in relation to the kymographic method of recording?
6. How may the cathode-ray oscilloscope be used to record the electrical activity of a cell?

Additional Reading

Brachet, J., and Mirsky, A. E., eds., "The Cell," vol. 1. Academic Press, New York, N. Y., 1958.

McManus, J. F. A., and Mowry, R. W., "Staining Methods Histologic and Histochemical," Paul B. Hoeber, Inc., New York, N. Y., 1960.

Price, W. J., "Nuclear Radiation Detection," McGraw-Hill Book Co., Inc., New York, N. Y., 1958.

Randall, J. E., "Elements of Biophysics," Yearbook Publishers, Chicago, Ill., 1958.

Umbreit, W. W., Burris, R. H., and Stauffer, J. F., "Manometric Tech-
niques and Related Methods for the Study of Tissue Metabolism," rev.
ed., Burgess Publishing Co., Minneapolis, Minn., 1957.
Willard, H. H., Merritt, L. L., and Dean, J. A., "Instrumental Methods
of Analysis," D. Van Nostrand Co., Inc., New York, N. Y., 1951.

INDEX

Illustrations are indexed with page numbers in boldface followed by the word **illus.;** tables, in boldface followed by **table.**

Date Due

NO 9 '62	JAN 4 1874	JAN 4 1974
NO 15 '62	OCT 30	NOV 22 1975
DE 17 '62	OCT 30	
DE 17 '62	APR 30 1978	APR 17 1978
JE 20 '63	MAR 15 1979	MAR 14 1979
OC 4 '63		
NO 1 '63		
NO 28 '63		
FE 15 '64		
RESERVE		
MAR 8 2 1966		
MAR 24 1966		
JUL 18 1965		
FEB 21 1967		
FEB 20 1969 FEB 1 9 1969		
JAN 15 1971		
DEC 2 1971 DE 2 71		
JUN 9 1972 JUN 2 1972		